Light Dosage Optimization by Data-driven and Dynamic Modeling in Blue Light Therapies

Wang, T.

DOI
10.4233/uuid:d3a84a3a-ba8d-4e1e-b6f7-4d3d2d7ac0c1

Publication date
2022

Document Version
Final published version

Citation (APA)
Wang, T. (2022). *Light Dosage Optimization by Data-driven and Dynamic Modeling: in Blue Light Therapies*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:d3a84a3a-ba8d-4e1e-b6f7-4d3d2d7ac0c1

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Light Dosage Optimization by Data-driven and Dynamic Modeling

in Blue Light Therapies

LIGHT DOSAGE OPTIMIZATION BY DATA-DRIVEN AND DYNAMIC MODELING

IN BLUE LIGHT THERAPIES

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University ofTechnology
by the authority of the Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Monday 27 June 2022 at 12:30 o'clock

by

Tianfeng WANG

Master of Science in Mechanical Engineering,
State university of New York at Buffalo,
born in Changchun, China.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. G.Q. Zhang,	Delft University of Technology, promotor
Prof. dr. J.F. Dong,	Suzhou Institute of Biomedical Engineering and Technology, copromotor
Independent members:	
Prof. dr. ir. W.D. van Driel,	Delft University of Technology
Prof. dr. R.A.M. Fouchier,	Erasmus University Medical Center
Prof. dr. D.A. Pijnappels,	Leiden University Medical Center
Prof. dr. Y.G. Tang,	Suzhou Institute of Biomedical Engineering and Technology
Prof. dr. ir. W.L. IJzerman,	Eindhoven University of Technology
Reserve members:	
Prof. dr. ir. P.M. Sarro,	Delft University of Technology

Keywords: Low-Level light therapy, Nonlinear dynamics, Mathematical model, Reactive oxygen species, Optimization algorithm

Printed by: Ipskamp Printing

Front & Back: Tianfeng Wang and GL-Never-Exert

ISBN 978-94-6421-784-1

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

Better to run than curse the road.

Clinkz, the Bone Fletcher

CONTENTS

SUMMARY

The therapeutic properties of light have been known for thousands of years, but photodynamic therapy (PDT) was only developed in the last century. Currently, PDT is in clinical trials for the oncology-the treatment of head and neck, brain, lung, pancreas, abdominal cavity, breast, prostate, and skin cancer. Advantages of the light-based therapies include rapid action and avoidance of drug resistance. The underlying mechanism of PDT is that the photosensitizers (PS) transform from their ground state (singlet state) into a relatively long-lived electronically excited state (triplet state) by the absorbing the photo energy, which, in turn, produces highly toxic reactive oxygen species (ROS) in cells. One difficulty of PDT is the administration of PS. Since the PS does not naturally exist, PDT relies on the exogenous PS which is administered by intravenous injection or topical application to the skin. This makes three disadvantages: first, the PS needs to be approved before it can be applied to patients; second, the adverse reaction of importing the exogenous PS can not be eliminated even if it is approved; and third, for the tumor under deep tissue, it is hard to import the PS to incidence area.

Similar to PDT, the antimicrobial blue light (ABL) only relies on the endogenous PS (flavin and porphyrin molecules) to inactivate the microbes, which is safer to use. However, as it is named, ABL can be only used for treating diseases whose pathogen is microbes, but the tumor. The most common application of ABL is treating various microbial superficial infections, e.g., skin or membranes. Traditionally, topical antimycotic/antibiotic drugs and more convenient oral azole agents are the main treatments for microbial infections. However, most pathogens have shown increased resistance to these drugs. Especially, the most famous one, methicillin-resistant *Staphylococcus aureus* (MRSA), was called "superbug" which has evolved resistance to most antibiotic drugs. Fortunately, ABL was proved to be effective in the inactivation of most pathogenic microbes, including MRSA, *Candida Albicans*, *Escherichia coli*. Further studies show that the inactivation effect did not significantly decrease after repeated ABL irradiation, which demonstrates the avoidance of resistance to ABL.

Aimed at the prediction of the light treatment outcome, some first-principle models have been proposed for estimating the treating dosages based on the mechanism of PDT and ABL. For instance, modeling of the dynamic changes in ROS concentrations has been well investigated, which are highly nonlinear models. These first-principle models precisely fit the dynamic changes in ROS concentrations. However, a dynamic model for the ROS concentrations to treat fungal infections has not been established. This can be attributed to the difficulty of using a first principle model, i.e., the parameters are related to the PS characteristics, which are determined experimentally. In contrast, although ABL is believed to be caused by the PS that naturally exists in microbial cells, whose types and amounts are usually unknown, first-principle modeling becomes even more challenging than modeling PDT.

Generally, to handle the challenges in modeling by first principles, system identifi-

cation methods (SIM) that estimate models from the data measured from complex dynamic processes have been well developed in control theory literature. Applying SIMs in modeling biological and biomedical systems is even more appealing than their traditional applications in mechanical, electrical systems, and industrial processes. This is because biological and biomedical systems possess even higher levels of nonlinearity, coupling, and chaos than traditional processes, and are hence more difficult to model with first principles. In recent decades, the applications of SIM methods to identify biological and biomedical systems, which are usually highly nonlinear, coupled, and chaotic, have also been witnessed.

Despite all the aforementioned efforts in biomedical system identification, identifying a nonlinear dynamic model for anti-fungal blue light therapy has not yet been targeted in the literature. In this thesis, we measure the viability, the total amount of ROS, and singlet oxygen (the main product of the type-II petrochemical reaction) in both pathogens and human host cells during the blue light irradiation 2. By analyzing these results, the optimized ABL wavelength and dosage were estimated. Furthermore, we have developed three independent models aimed at estimating the ABL dosages. The first one is a liner integrator model, which is based on a first-order dynamic equation 3. The second one is a kernel-based NARX model with an optimized kernel selection method 4. The third one is a closed-formed equation, which was derived from a traditional microscopic singlet oxygen model 6. All of the three models offer opportunities for determining the light dosages in treating fungal infection diseases by ABL. To the best of our knowledge, it is also the first attempt of applying control theory to analyze the photochemical reaction dynamics of light therapies.

SAMENVATTING

De therapeutische eigenschappen van licht zijn al duizenden jaren bekend, maar fotodynamische therapie (PDT) werd pas in de vorige eeuw ontwikkeld. Momenteel bevindt PDT zich in klinische onderzoeken voor de oncologie - de behandeling van hoofd en nek, hersenen, longen, pancreas, buik holte-, borst-, prostaat- en huidkanker. Voordelen van de op licht gebaseerde therapieën zijn onder meer een snelle actie en het vermijden van resistentie tegen geneesmiddelen. Het onderliggende mechanisme van PDT is dat de fotosensitizers (PS) van hun grondtoestand (singlettoestand) transformeren in een relatief langlevende elektronisch aangeslagen toestand (triplettoestand) door het absorberen van de foto-energie, die op zijn beurt zeer giftige reactieve zuurstofsoorten (ROS) in cellen. Een moeilijkheid van PDT is de toediening van PS. Aangezien de PS van nature niet bestaat, vertrouwt PDT op de exogene PS die wordt toegediend via intraveneuze injectie of plaatselijke toepassing op de huid. Dit heeft drie nadelen: ten eerste moet de PS worden goedgekeurd voordat deze op patiënten kan worden toegepast; ten tweede kan de negatieve reactie van het importeren van de exogene PS niet worden geëlimineerd, zelfs niet als deze is goedgekeurd; en ten derde, voor de tumor onder diep weefsel, is het moeilijk om de PS naar het incidentiegebied te importeren.

Net als bij PDT, vertrouwt het antimicrobiële blauwe licht (ABL) alleen op de endogene PS (flavine- en porfyrinemoleculen) om de microben te inactiveren, wat veiliger is om te gebruiken. Zoals het wordt genoemd, kan ABL echter alleen worden gebruikt voor de behandeling van ziekten waarvan de ziekteverwekker microben zijn, maar de tumor. De meest voorkomende toepassing van ABL is de behandeling van verschillende microbiële oppervlakkige infecties, bijvoorbeeld huid of vliezen. Traditioneel zijn lokale antimycotica/antibiotica en gemakkelijkere orale azoolmiddelen de belangrijkste behandelingen voor microbiële infecties. De meeste pathogenen hebben echter een verhoogde resistentie tegen deze medicijnen laten zien. Vooral de meest bekende, methicilline-resistente *Staphylococcus aureus* (MRSA), heette "superbug", die resistentie heeft ontwikkeld tegen de meeste antibiotica. Gelukkig bleek ABL effectief te zijn bij het inactiveren van de meeste pathogene microben, waaronder MRSA, *Candida Albicans*, *Escherichia coli*. Verdere studies tonen aan dat het inactiveringseffect niet significant afnam na herhaalde ABL-bestraling, wat het vermijden van resistentie tegen ABL aantoont.

Gericht op de voorspelling van het resultaat van de lichtbehandeling, zijn enkele eerste-principemodellen voorgesteld voor het schatten van de behandelingsdoseringen op basis van het mechanisme van PDT en ABL. Het modelleren van de dynamische veranderingen in ROS concentraties is bijvoorbeeld goed onderzocht, wat zeer niet-lineaire modellen zijn. Deze eerste-principemodellen passen precies bij de dynamische veranderingen in ROS-concentraties. Er is echter geen dynamisch model vastgesteld voor de ROS-concentraties om schimmelinfecties te behandelen. Dit kan worden toegeschreven aan de moeilijkheid om een eerste principemodel te gebruiken, d.w.z. de parameters zijn gerelateerd aan de PS-kenmerken, die experimenteel worden bepaald. Daarente-

gen, hoewel wordt aangenomen dat ABL wordt veroorzaakt door de PS die van nature voorkomt in microbiële cellen, waarvan de typen en hoeveelheden meestal onbekend zijn, wordt het modelleren van het eerste principe zelfs nog uitdagender dan het modelleren van PDT.

Over het algemeen zijn systeemidentificatiemethoden (SIM) die modellen schatten op basis van de gemeten gegevens van complexe dynamische processen, goed ontwikkeld in de literatuur over regeltheorie om de uitdagingen bij modellering op basis van eerste principes aan te pakken. Het toepassen van SIM's bij het modelleren van biologische en biomedische systemen is nog aantrekkelijker dan hun traditionele toepassingen in mechanische, elektrische systemen en industriële processen. Dit komt omdat biologische en biomedische systemen nog hogere niveaus van niet-lineariteit, koppeling en chaos hebben dan traditionele processen, en daarom moeilijker te modelleren zijn met de eerste principes. In de afgelopen decennia zijn ook de toepassingen van SIM-methoden waargenomen om biologische en biomedische systemen te identificeren, die meestal zeer niet-lineair, gekoppeld en chaotisch zijn.

Ondanks alle bovengenoemde inspanningen op het gebied van identificatie van biomedische systemen, is het identificeren van een niet-lineair dynamisch model voor antischimmeltherapie met blauw licht nog niet het doelwit van de literatuur. In dit proefschrift meten we de levensvatbaarheid, de totale hoeveelheid ROS en singlet zuurstof (het belangrijkste product van de type-II petrochemische reactie) in zowel pathogenen als menselijke gastheercellen tijdens de bestraling met blauw licht 2. Door deze resultaten te analyseren, werden de geoptimaliseerde ABL-golflengte en -dosering geschat. Verder hebben we drie onafhankelijke modellen ontwikkeld om de ABL-doseringen te schatten. De eerste is een liner integrator-model, dat is gebaseerd op een eerste-orde dynamische vergelijking 3. De tweede is een op kernels gebaseerd NARX-model met een geoptimaliseerde kernelselectiemethode 4. De derde is een gesloten formule, die is afgeleid van een traditioneel microscopisch singlet-zuurstofmodel 6. Alle drie de modellen bieden mogelijkheden voor het bepalen van de lichtdoseringen bij de behandeling van schimmelinfecties door ABL. Voor zover wij weten, is het ook de eerste poging om controletheorie toe te passen om de fotochemische reactiedynamiek van lichttherapieën te analyseren.

PREFACE

Dear reader,

This thesis is the result of four years of Ph.D work. During that time, I have been immersed in an environment of highly talented people. After so many hours in this community, I have become accustomed to a particular jargon and way of writing – maybe accustomed too much. For that reason, let me give a brief 'reading guide'

First, most researchers do not read scientific articles like a novel: from front to back. Instead, they read the title and summary, then skim through the paper and look at the figures. If the article still looks interesting, then they read the introduction, and potentially also the conclusion. By doing so, the reader establishes a 'frame' in which the rest of the information of the article can be put (You could compare it to a wardrobe that contains many drawers: it is neater to first have the wardrobe and then fill it with clothes, then to first have a pile of clothes on the ground, after which you will build the wardrobe). With this frame, it is much easier to read the rest of the paper – this time from front to back.

I propose that you, as reader of this dissertation, do exactly the same, at least at the start: first read the title and the summary, then proceed to the introduction. At the end of the introduction I will give an overview of the different chapters, which hopefully gives you an idea of what you would like to read next. Like most theses in the hard sciences, the core of this dissertation consists of chapters which are modified versions of published articles. Although these were written for an expert audience, please do not feel held back to read them by reading title and summary, skim the figures, etc.

I believe the use of jargon can hardly ever be completely avoided. However, I have noticed that many around me are increasingly often using field-specific terminology in everyday conversations, often without noticing, and it would be naive to think I am any different. For that reason, I have tried to compile a list of words that might be useful to know when reading this thesis – already starting with the words in the title. If you find that the dissertation is completely incomprehensible after having read this list and the introductory chapters, then I welcome you to come over for a cup of tea (or something stronger) and a good chat. I am more than willing to explain this dissertation's content.

So here comes the list of words.

First of all, you will find that this dissertation is written in 'we' form (first person plural). Partially I decided to do so because this is common in the field. More importantly, a lot of the work in the core chapters of this dissertation was done in collaboration with others. For those chapters, it would therefore be not only impolite to say that 'I did such and such', but simply incorrect.

Next, the word 'photosensitizer (PS)' is the definition of a group of molecules which absorb light (hv) and transfer the energy from the incident light into another nearby molecule. You may still confusing, do not worry. You can just understand PS as something can react with light and in turn lead to the treatment.

The most frequently used words in this thesis maybe 'PDT' and 'ABL', they are the abbreviations of 'photodynamic therapy' and 'anti-microbial blue light' therapy respectively. Both of them are the including in the definition of light therapies. However, PDT is a more general definition which can be applied with any light, while ABL only relies on blue light. Another difference is the target diseases, PDT was mainly used for cancer treatment. In contrast, ABL was investigated on superficial diseases due to lacking of penetration. The mechanism of PDT and ABL were similar, both rely on the PS, i.e PDT relies on exogenous PS and ABL relies on endogenous PS.

The italic *in vivo* means 'within the living' in Latin. In biological readings, this words indicates the experiments were performed on animals. In contrast, the word *in vitro* means 'in glass', which indicates the studies were performed with microorganisms, cells, or biological molecules outside their normal biological context.

The abbreviations *i.e.* and *e.g.*, or *id est* and *exempli gratia*, mean 'that is' and 'for example', respectively.

Although this list is far from complete, I hope it helps in reading this thesis a bit more easily. Let me repeat that I am more than willing to come talk to you to explain what I have been doing for the past four years.

Tianfeng Wang
May 2022

1

INTRODUCTION

1

1.1. PHOTODYNAMIC THERAPY

Thousands of years ago, the therapeutic properties of light have been realized. However, photodynamic therapy (PDT) was only developed in the last century. Currently, PDT is in clinical trials for oncology, i.e., the treatment of head and neck, brain, pancreas, lung, abdominal cavity, breast, prostate, and skin cancers.

Light has been used in disease treatment for thousands of years [1, 2]. Ancient Indian, Egyptian, and Chinese people discovered the usage of light to treat various diseases, including psoriasis, rickets, vitiligo, and skin cancer [3]. At the end of the 19th century, 'phototherapy' was developed to treat diseases by Niels Finsen. He found that red light irradiation can prevent the discharge and formation of smallpox pustules and can be used as an alternative treatment for this disease [4]. Then, he used solar ultraviolet rays to treat skin tuberculosis. So here is the beginning of modern light therapies (see in the timeline in Fig. 1.1).

Hundred years ago, it was observed that light irradiation with certain chemicals can lead to cell death. In 1900, certain wavelengths were lethal to infusoria (including a species of Paramecium) with the presence of acridine was found by German medical student Oscar Raab [5]. In the same year, a French neurologist named J. Prime discovered dermatitis in sun-exposed areas in patients with epilepsy taking oral eosin treatment [6]. In 1903, Herman Von Tappeiner and A. Jesionek used the combination of white light and topical application of eosin to treat skin tumors [7]; this phenomenon was finally described as 'PHOTODYNAMIC ACTION' [8].

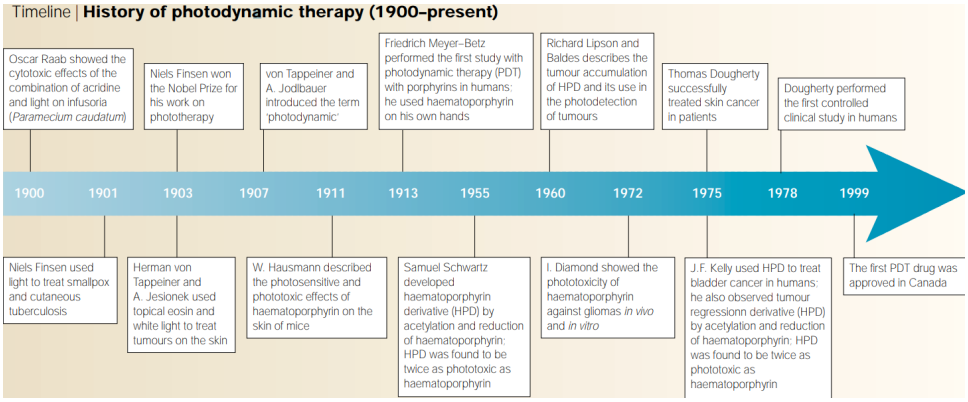


Figure 1.1: History of photodynamic therapy (since 1990) [9]

Experiments for testing the combination of reagents and light led to modern photodynamic therapy (PDT). Two independent non-toxic components combine in an oxygen-dependent manner to induce cellular and tissue damage in PDT (as shown in Fig. 1.2). The first component is a photosensitizer (PS), which is a kind of molecule that is localized to target cells and tissues. The second component involves the management of specific wavelengths of light that excite the PS. The PS is able to transfer energy from light to molecular oxygen, then lead to the generation of reactive oxygen species (ROS). These reactions occur in the direct area where the light absorbs the photosensitizer. Therefore,

the biological response to the photosensitizer is only activated in the specific tissue area exposed to light. Other photochemical reactions are not using oxygen as an intermediate. For example, photoaddition to DNA has also been developed in [10]. These reactions are called 'photochemotherapy'. One photochemotherapeutic, called 'psoralens', has been combined with UVA to treat psoriasis, and vitiligo and to enhance immunotherapy [11].

Porphyrins were the most common PS and have been widely studied since mid-nineteenth century. These compounds contain a porphin structure—four pyrrole rings connected in a cyclic structure by a methyl bridge—and a side chain that is usually a metal. For example, the combination of iron and porphine structure forms heme. W. Hausmann used these reagents for the first study [12]. He treated Paramecium and red blood cells with hematoporphyrin and light, and reported that this combination killed these cells. In addition, he also reported skin reactions in mice exposed to light after hematoporphyrin administration. In 1913, the German scientist Friedrich Meyer–Betz was the first one to treat humans with porphyrins, i.e., using 200 mg of haematoporphyrin on his skin for testing, which lead to swelling and pain in light-exposed areas [13].

In the 1960s, Richard Lipson and his colleagues start to investigate the modern PDT at the Mayo Clinic [14, 15]. In 1996, Samuel Schwartz developed a compound called 'haematoporphyrin derivative' (HPD) [16]. To prepare this derivative, crude haematoporphyrin was treated with acetic and sulphuric acids, filtered, and then neutralized with sodium acetate. The precipitate was then resolved in saline to produce HPD. Lipson and E.J. Baldes then showed that HPD can localize to tumours, and emit fluorescence. Because this derivative can act at much smaller doses than crude haematoporphyrin, it can be used as a diagnostic tool [17]. The mechanisms by which photosensitizers such as HPD selectively accumulate in tumors are complex and not fully understood. It is probably because of the high vascular permeability of the agents, as well as their affinity for proliferating endothelium and the lack of lymphatic drainage in tumors [18, 19].

In 1903, the first experiments on the therapeutic application of PDT to patients with cancer were carried out by Von Tappeiner and Jesionek. In 1972, I. Diamond and colleagues hypothesized that the combination of the tumor-localizing and tumor-phototoxic properties of porphyrins might be exploited to kill cancer cells [20]. In *vivo*, studies revealed that PDT inactivated the growth of gliomas that were implanted in rats. Tumor growth was suppressed for 10–20 days, but finally deeper regions of the tumors began growing again. In 1975 Thomas Dougherty and co-workers reported that administration of HPD and red light completely eliminated mammary tumor growth in mice, this is a milestone breakthrough. [21]. In the same year, J.F. Kelly and co-workers reported that light activation of HPD can also eliminate bladder carcinoma in mice [22].

In 1976, Kelly and co-workers carried out the first human trials with HPD. More specifically, five patients were diagnosed with cancer using HPD. It was also used to treat one patient with recurrent bladder carcinoma who had failed transurethral resections, radiotherapy, and chemotherapy. In this patient, HPD inactivated tumour growth, and tumour necrosis was observed in the areas that received PDT. In a second study by Dougherty et al., 25 patients with a total of 113 primary or secondary skin tumours were treated with HPD. A complete response was observed in 98 tumours, a partial response

1

was observed in 13 tumours and 2 tumours were found to be treatment-resistant [23]. With these preliminary successes, Y. Hayata and colleagues used PDT to treat obstructing lung tumours [24]. Bronchoscopic analysis revealed tumor growth delay in most patients, but only one out of fourteen patients were completely cured.

Then, J.S. McCaughan used PDT to treat oesophageal cancer in patients in 1984, [25]. In the same year, O.J. Balchum and colleagues used PDT to treat patients with lung cancer [26]. One year later, Y. Hayata and colleagues used PDT to treat gastric carcinoma on patients[27]. All of these studies showed advantageous responses in early-stage patients, so PDT was recommended for patients with early-stage cancers that were inoperable. Patients with breast cancer [28], gynaecological tumours [29], intraocular tumours, brain tumours [30], head and neck tumours [31], colorectal cancer, cutaneous malignancies [32], intraperitoneal tumours, mesothelioma, cholangiocarcinoma [33] and pancreatic cancer [34] were subsequently treated with PDT. However, PDT has only shown limited success in further studies, due to issues of the specificity and potency of PS. Another confusing factor is that PDT has been tested mainly in patients with advanced-stage diseases that are obstinate to other treatments. In such cases, a local effect cannot usually significantly alter the outcome of systemic disease. More selective and potent sensitizers have been developed, and are now under investigation in clinical trials. With these newly developed sensitizers, better localization methods, and improved protocols and equipment, the efficacy of PDT might be enhanced [35].

1.2. MECHANISM OF ACTION

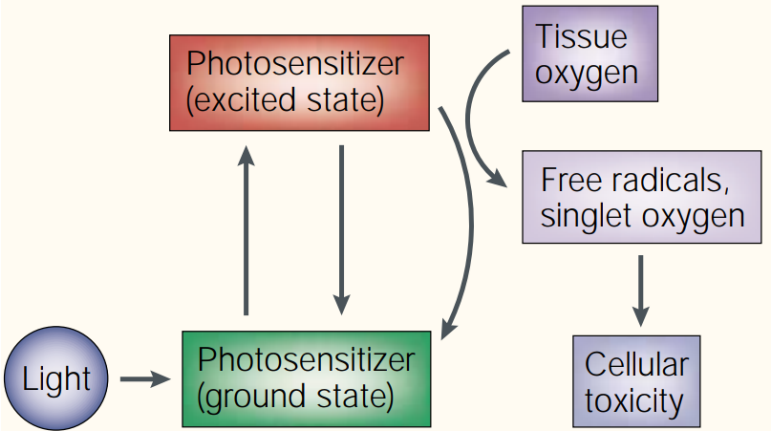


Figure 1.2: Mechanism of action of photodynamic therapy (PDT). PDT requires three elements: light, a photosensitizer and oxygen. When the photosensitizer is exposed to specific wavelengths of light, it becomes activated from the ground to an excited state. As it returns to the ground state, it releases energy, which is transferred to oxygen to generate reactive oxygen species (ROS), such as singlet oxygen and free radicals. These ROS mediate cellular toxicity [9].

One advantage of PDT is that the PS can be administered by various techniques, i.e., intravenous injection and topical application to the skin. By absorbing the energy of light, the PS is transformed from its ground state (singlet state) into a relatively long-

lived electronically excited state (triplet state) via a short-lived excited singlet state [36]. As shown in Fig. 1.3, two types of petrochemical reactions can lead to the generation of the excited state. First, it can react directly with a substrate, such as a cell membrane or a molecule, and transfer a hydrogen atom (electron) to form radicals. These radicals interact with oxygen to produce oxygenated products (type I reaction). Alternatively, the triplet can transfer its energy directly to oxygen, to form singlet oxygen — a highly ROS (type II reaction). Because the effects of almost all PDT drugs are oxygen dependent, photosensitization typically does not occur in anoxic areas of tissue. It was reported that the PDT effects were abolished at hypoxia *in vivo*, inducted by clamping the tissue [37].

Both type I and type II reactions occur simultaneously, and various conditions lead to the change of the ratio between these processes, i.e., the type of PS used, the concentrations of substrate and oxygen, as well as the binding affinity of the PS for the substrate. Because of the high reactivity and short half-life of the ROS, only cells that are proximal to the area of the ROS production (areas of PS localization) are directly affected by PDT [38]. The half-life time of singlet oxygen in biological systems is less than $0.04\mu s$. Therefore, the reaction radius of singlet oxygen is also very limited, i.g., less than $0.02\mu m$. The extent of photodamage and cytotoxicity is multifactorial and depends on the type of PS, its extracellular and intracellular localization, the total dose administered, the total light dosage, light fluence rate, oxygen availability, and the time between the administration of the drug and light exposure. All of these factors are correlated and therefore interdependent.

1.3. ANTI-FUNGAL BLUE LIGHT

At present, antibiotic and antimycotic drugs are the main treatments for microbial infection. However, the common pathogens, such as Candidal species, *Staphylococcus aureus*, and *Escherichia coli*, are showing an increasing resistance to these drugs. Therefore, it is crucial to find alternative treatments. As a non-antibiotic approach, light-based antimicrobial therapies, including antimicrobial photodynamic therapy (aPDT) and ultraviolet-C (UVC) irradiation, have been extensively investigated as alternative therapeutics for infectious diseases, especially for topical localized infections [39, 40]. Advantages of light-based therapies include rapid action and equal inactivation effectiveness regardless of drug resistance [41, 42]. However, one major disadvantage of aPDT, as a two-part (PS and light) combination approach, is the challenge of introducing exogenous PS into certain pathogens and less than perfect selectivity for pathogens over host cells [43]. The use of UVC, on the other hand, has limitations due to its detrimental effects on host cells [40]. An innovative light-based antimicrobial approach, antimicrobial blue light (ABL), has attracted increasing attention due to its intrinsic ability to inactivate pathogens without the involvement of exogenous PS [44].

The mechanism of action of ABL is still not fully understood. A common hypothesis is that ABL is similar to PDT, which can excite naturally occurring endogenous PS in the cells of pathogens and subsequently leads to the production of cytotoxic oxidative species [44, 45]. Due to the type and content of endogenous PS in different cells, their susceptibility to ABL can also be different. Fungal PS content is normally higher than that of human cells. Thus, fungi are more susceptible to ABL and, hence, ABL has been widely investigated for treating fungal infections. For instance, the inactivation rate of

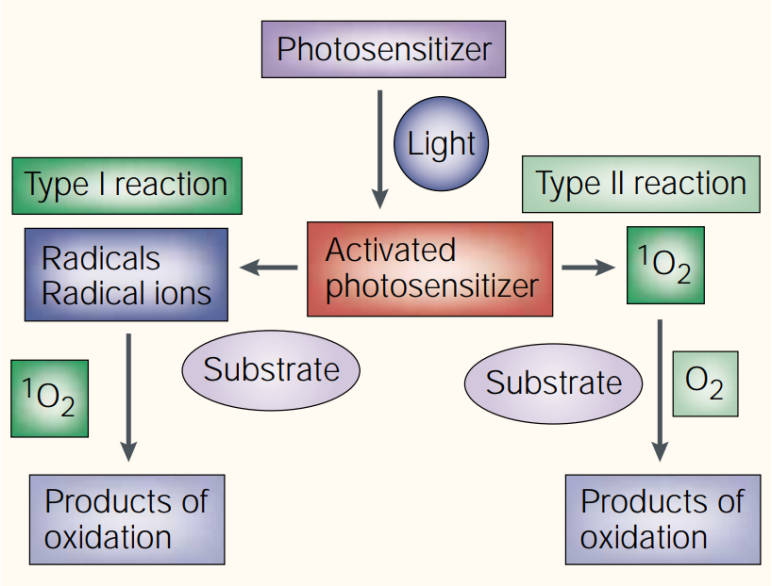


Figure 1.3: Type I and type II reaction in photodynamic therapy (PDT). There are two types of reactions during PDT. Following the absorption of light, the PS is transformed from its ground state into an excited state. The activated PS can undergo two kinds of reactions. First, it can react directly either with the substrate, such as the cell membrane or a molecule, transferring a hydrogen atom to form radicals. The radicals interact with oxygen to produce oxygenated products ($^1O_2$) (type I reaction). Alternatively, the activated PS can transfer its energy directly to oxygen, to form singlet oxygen ($^1O_2$) — a highly reactive oxygen species. These species oxidize various substrates (type II reaction).

C. albicans by ABL was 42-fold faster than human keratinocytes [46]. A dynamic model was developed based on the viability of *C. albicans* and vaginal epithelial (V. E.) cells during ABL irradiation, which demonstrated that the shorter ABL wavelength around 410nm achieved a higher anti-fungal effect than 450nm [47]. The safety of ABL to the V. E. cells was investigated in [48]. The blue light at 405nm preferentially inactivated the pathogen over the V. E. cells. Furthermore, no genotoxicity of blue light to the V. E. cells was observed at the dosage for inactivating the pathogen.

However, using ABL for treating actual infections was not established yet. The ABL investigation was still confined to *in vitro* efficacy studies in recent decade [49]. To the best of our knowledge, only a few *in vivo* studies report the efficacy of ABL for infections [50–52]. More specifically, it has been demonstrated that ABL at 415nm significantly reduced the bacterial infection in mouse wounds or burns which caused by both Gram-positive and Gram-negative bacteria [50, 52], and successfully rescue the mouse with lethal P. aeruginosa infection [51].

1.4. MODELING APPROACHES

Based on the two types of petrochemical reactions, some first-principle models have been proposed [53–55], which are highly nonlinear models. Many factors need to be con-

sidered in these models, e.g., photosensitizer and its photochemical parameters, oxygen consumption and supply rates, and irradiation conditions. Therefore, these kinds of models are usually very complex and the parameters included are hard to determine. For instance, up to 21 parameters are required to describe the process related to the PS [54]. Although ABL is believed to be caused by the endogenous PS, whose types and amounts are usually unknown, first-principle modeling becomes even more challenging than modeling PDT. To avoid the aforementioned problems, simplification of these models or data-driven modeling are two research aims in this field.

Data-driven modeling estimates models from the data measured from complex dynamic processes, where the system identification methods were well developed in control theory literature [56]. Applying SIMs in modeling biological and biomedical systems is even more appealing than their traditional applications in mechanical, electrical systems, and industrial processes. This is because biological and biomedical systems possess even higher levels of nonlinearity, coupling, and chaos than traditional processes, and are hence more difficult to model with first principles. In the recent decades, SIM methods were used to identify nonlinear, coupled, and chaotic biological and biomedical systems [57–59]. The data-driven models are usually "black box" models, whose internal laws are rarely known. Therefore, after the model was established, it is necessary to explain the model parameters with physical meaning. However, we could not always find a proper meaning for all parameters. This could be one limitation of the data-driven models.

On the other hand, the simplification based on the complex first-principle models does not have the same problem. Every parameter was given physical meanings as the model was established. With some reasonable assumptions, the macroscopic singlet oxygen model [55] was simplified into three differential equations in [60] as follows.

$$\frac{d[S_0]}{dt} = -\xi\sigma\frac{\phi([S_0]+\delta)[^3O_2]}{[^3O_2]+\beta}[S_0] \tag{1.1a}$$

$$\frac{d[^3O_2]}{dt} = -\xi\frac{\phi[S_0]\,[^3O_2]}{[^3O_2]+\beta} + g\cdot(1-\frac{[^3O_2]}{[^3O_2]_0}) \tag{1.1b}$$

$$\frac{d[^1O_2]}{dt} = f\cdot\xi\frac{\phi[S_0][^3O_2]}{[^3O_2]+\beta} \tag{1.1c}$$

The original 21 parameters were merged or neglected. However, still seven parameters remained. More importantly, the equation's type is kept in differential type, which brings more computational difficulty.

1.5. AIM OF THE THESIS

Since PDT and ABL are relatively new therapeutic methods for cancer and nonmalignant disease treatment in recent hundred years. The clinical and *in vivo* studies were relatively less, especially for ABL therapies. To the best of our knowledge, the ABL past clinical and *in vivo* studies do not have specific dosage guidance. The experimental procedure was starting with a small light dosage and watching the treatment outcome. Usually, biologists focus more on the experiments and data than on the analysis of the data. Until now, no nonlinear model based on experimental data has been established. Fur-

thermore, no study has analyzed the effects and safety of ABL therapy based on a dynamic ROS model. It is the purpose of this thesis to optimize the light dosages in ABL, through both experiments and modeling.

More specifically, we measured the viability, the total amount of ROS, and singlet oxygen (the main product of the type-II petrochemical reaction) in both pathogens and human host cells during the blue light irradiation in Chapter 2. By analyzing these results, the optimized ABL wavelength and dosage were estimated. Furthermore, we have developed three independent models aimed at estimating the ABL dosages. The first one is a liner integrator model, which is based on a first-order dynamic equation (shown in Chapter 3). The second one is a kernel-based NARX model with an optimized kernel selection method (shown in Chapter 4). The third one is a closed-formed equation, which was derived from a traditional microscopic singlet oxygen model (shown in Chapter 6). All of the three models offer opportunities for determining the light dosages in treating fungal infection diseases by ABL. To the best of our knowledge, it is also the first attempt of applying control theory to analyze the photochemical reaction dynamics of light therapies.

REFERENCES

[1] M. Daniell and J. Hill, *A history of photodynamic therapy,* Australian and New Zealand Journal of Surgery **61**, 340 (1991).

[2] *The history of photodetection and photodynamic therapy¶,* .

[3] R. V. Bensasson, *Primary photo-processes in biology and medicine,* **85** (2013).

[4] N. Finsen and J. Phototherapy, *Sequira (transl.),* Edward Arnold, London , 79 (1901).

[5] H. Von Tappeiner, *Uber die wirkung fluoreszierender stoffe auf infusorien nach versuchen von o,* Raab Muench Med Wochenschr **47** (1900).

[6] C. S. Huettner, P. Zhang, R. A. Van Etten, and D. G. Tenen, *Reversibility of acute b-cell leukaemia induced by bcr–abl1,* Nature genetics **24**, 57 (2000).

[7] H. Von Tappeiner, *Therapeutische versuche mit fluoreszierenden stoffen.* Munch Med Wochenschr **1**, 2042 (1903).

[8] H. Von Tappeiner and A. Jodlbauer, *Die sensibilisierende wirkung fluorieszierender substanzer. gesammte untersuchungen uber die photodynamische erscheinung,* FCW Vogel, Leipzig (1907).

[9] D. E. Dolmans, D. Fukumura, and R. K. Jain, *Photodynamic therapy for cancer,* Nature reviews cancer **3**, 380 (2003).

[10] J. H. Epstein, *Phototoxicity and photoallergy. Seminars in cutaneous medicine and surgery,* **18**, 274 (1999).

[11] J. H. Epstein, *Phototherapy and photochemotherapy,* (1990).

[12] W. Hausmann, *Über die sensibilisierende Wirkung des Hämatoporphyrins* (1910).

[13] F. Meyer-Betz *et al., Untersuchungen über die biologische (photodynamische) wirkung des hämatoporphyrins und anderer derivate des blut-und gallenfarbstoffs,* Dtsch. Arch. Klin. Med **112**, 0366 (1913).

[14] R. L. LIPSON and E. J. BALDES, *The photodynamic properties of a particular hematoporphyrin derivative,* Archives of dermatology **82**, 508 (1960).

[15] R. L. Lipson, E. J. Baldes, and A. M. Olsen, *The use of a derivative of hematoporphyrin in tumor detection,* Journal of the National Cancer Institute **26**, 1 (1961).

[16] S. Schwartz, K. Absolon, and H. Vermund, *Some relationships of porphyrins, x-rays and tumors,* Univ. Minn. Med. Bull **27** (1955).

[17] R. L. Lipson, E. J. Baldes, and A. M. Olsen, *Hematoporphyrin derivative: a new aid for endoscopic detection of malignant disease,* The Journal of thoracic and cardiovascular surgery **42**, 623 (1961).

[18] T. J. Dougherty, C. J. Gomer, B. W. Henderson, G. Jori, D. Kessel, M. Korbelik, J. Moan, and Q. Peng, *Photodynamic therapy,* JNCI: Journal of the national cancer institute **90**, 889 (1998).

[19] D. W. Kufe, J. F. Holland, and E. Frei, *Cancer medicine 6*, Vol. 2 (BC Decker, 2003).

[20] I. Diamond, A. Mcdonagh, C. Wilson, S. Granelli, S. Nielsen, and R. Jaenicke, *Photodynamic therapy of malignant tumours,* The Lancet **300**, 1175 (1972).

[21] T. J. Dougherty, G. Grindey, R. Fiel, K. R. Weishaupt, and D. Boyle, *Photoradiation therapy. ii. cure of animal tumors with hematoporphyrin and light,* Journal of the National Cancer Institute **55**, 115 (1975).

[22] J. Kelly, M. Snell, and M. Berenbaum, *Photodynamic destruction of human bladder carcinoma,* British journal of cancer **31**, 237 (1975).

[23] T. J. Dougherty, J. E. Kaufman, A. Goldfarb, K. R. Weishaupt, D. Boyle, and A. Mittleman, *Photoradiation therapy for the treatment of malignant tumors,* Cancer research **38**, 2628 (1978).

[24] Y. Hayata, H. Kato, C. Konaka, J. Ono, and N. Takizawa, *Hematoporphyrin derivative and laser photoradiation in the treatment of lung cancer,* Chest **81**, 269 (1982).

[25] J. S. McCaughan Jr, W. Hicks, L. Laufman, E. May, and R. Roach, *Palliation of esophageal malignancy with photoradiation therapy,* Cancer **54**, 2905 (1984).

[26] O. J. Balchum, D. R. Doiron, and G. C. Huth, *Photoradiation therapy of endobronchial lung cancers employing the photodynamic action of hematoporphvrin derivative,* Lasers in surgery and medicine **4**, 13 (1984).

[27] Y. Hayata, H. Kato, H. Okitsu, M. Kawaguchi, and C. Konaka, *Photodynamic therapy with hematoporphyrin derivative in cancer of the upper gastrointestinal tract,* in *Seminars in surgical oncology*, Vol. 1 (Wiley Online Library, 1985) pp. 1–11.

[28] T. J. Dougherty, G. Lawrence, J. H. Kaufman, D. Boyle, K. R. Weishaupt, and A. Goldfarb, *Photoradiation in the treatment of recurrent breast carcinoma,* Journal of the National Cancer Institute **62**, 231 (1979).

[29] B. G. Ward, C. PA, and M. MM, *The treatment of vaginal recurrences of gynecologic malignancy with phototherapy following hematoporphyrin derivative pretreatment,* (1982).

[30] I. M. Landau, B. Steen, and S. Seregard, *Photodynamic therapy for circumscribed choroidal haemangioma,* Acta Ophthalmologica Scandinavica **80**, 531 (2002).

[31] V. G. Schweitzer, *Photodynamic therapy for treatment of head and neck cancer,* Otolaryngology—Head and Neck Surgery **102**, 225 (1990).

[32] R. R. Allison, T. S. Mang, and B. D. Wilson, *Photodynamic therapy for the treatment of nonmelanomatous cutaneous malignancies.* in *Seminars in cutaneous medicine and surgery*, Vol. 17 (1998) pp. 153–163.

[33] M.-A. E. Ortner, J. Liebetruth, S. Schreiber, M. Hanft, U. Wruck, V. Fusco, J. M. Müller, H. Hörtnagl, and H. Lochs, *Photodynamic therapy of nonresectable cholangiocarcinoma,* Gastroenterology **114**, 536 (1998).

[34] S. Bown, A. Rogowska, D. Whitelaw, W. Lees, L. Lovat, P. Ripley, L. Jones, P. Wyld, A. Gillams, and A. Hatfield, *Photodynamic therapy for cancer of the pancreas,* Gut **50**, 549 (2002).

[35] G. McBride, *Studies expand potential uses of photodynamic therapy,* Journal of the National Cancer Institute **94**, 1740 (2002).

[36] B. W. Henderson and T. J. Dougherty, *How does photodynamic therapy work?* Photochemistry and photobiology **55**, 145 (1992).

[37] C. J. Gomer and N. J. Razum, *Acute skin response in albino mice following porphyrin photosensitization under oxic and anoxic conditions,* Photochemistry and photobiology **40**, 435 (1984).

[38] J. Moan and K. BERG, *The photodegradation of porphyrins in cells can be used to estimate the lifetime of singlet oxygen,* Photochemistry and photobiology **53**, 549 (1991).

[39] T. Dai, Y.-Y. Huang, and M. R. Hamblin, *Photodynamic therapy for localized infections—state of the art,* Photodiagnosis and photodynamic therapy **6**, 170 (2009).

[40] T. Dai, M. S. Vrahas, C. K. Murray, and M. R. Hamblin, *Ultraviolet c irradiation: an alternative antimicrobial approach to localized infections?* Expert review of anti-infective therapy **10**, 185 (2012).

[41] T. Maisch, *A new strategy to destroy antibiotic resistant microorganisms: antimicrobial photodynamic treatment,* Mini Rev. Med. Chem. **9**, 974 (2009).

[42] T. Conner-Kerr, P. Sullivan, J. Gaillard, M. Franklin, and R. Jones, *The effects of ultraviolet radiation on antibiotic-resistant bacteria in vitro.* Ostomy/wound management **44**, 50 (1998).

[43] M. Wainwright, *Photodynamic antimicrobial chemotherapy (pact).* The Journal of antimicrobial chemotherapy **42**, 13 (1998).

[44] R. Yin, T. Dai, P. Avci, A. E. S. Jorge, W. C. de Melo, D. Vecchio, Y.-Y. Huang, A. Gupta, and M. R. Hamblin, *Light based anti-infectives: ultraviolet c irradiation, photodynamic therapy, blue light, and beyond,* Current opinion in pharmacology **13**, 731 (2013).

[45] R. Yin, T. Dai, P. Avci, A. E. S. Jorge, W. C. de Melo, D. Vecchio, Y.-Y. Huang, A. Gupta, and M. R. Hamblin, *Light based anti-infectives: ultraviolet c irradiation, photodynamic therapy, blue light, and beyond,* Curr. Opin. Pharmacol. **13**, 731 (2013).

[46] Y. Zhang, Y. Zhu, J. Chen, Y. Wang, M. E. Sherwood, C. K. Murray, M. S. Vrahas, D. C. Hooper, M. R. Hamblin, and T. Dai, *Antimicrobial blue light inactivation of candida albicans: In vitro and in vivo studies,* Virulence **7**, 536 (2016).

[47] T. Wang, J. Dong, H. Yin, and G. Zhang, *Blue light therapy to treat candida vaginitis with comparisons of three wavelengths: an in vitro study,* Lasers Med. Sci. **35**, 1329 (2020).

[48] Y. Wang, R. Ferrer-Espada, Y. Baglo, X. S. Goh, K. D. Held, Y. H. Grad, Y. Gu, J. A. Gelfand, and T. Dai, *Photoinactivation of neisseria gonorrhoeae: a paradigm-changing approach for combating antibiotic-resistant gonococcal infection,* J. Infect.Dis. **220**, 873 (2019).

[49] C. S. Enwemeka, D. Williams, S. K. Enwemeka, S. Hollosi, and D. Yens, *Blue 470-nm light kills methicillin-resistant staphylococcus aureus (mrsa) in vitro,* Photomedicine and laser surgery **27**, 221 (2009).

[50] T. Dai, A. Gupta, Y.-Y. Huang, M. E. Sherwood, C. K. Murray, M. S. Vrahas, T. Kielian, and M. R. Hamblin, *Blue light eliminates community-acquired methicillin-resistant staphylococcus aureus in infected mouse skin abrasions,* Photomedicine and laser surgery **31**, 531 (2013).

[51] T. Dai, A. Gupta, Y.-Y. Huang, R. Yin, C. K. Murray, M. S. Vrahas, M. E. Sherwood, G. P. Tegos, and M. R. Hamblin, *Blue light rescues mice from potentially fatal pseudomonas aeruginosa burn infection: efficacy, safety, and mechanism of action,* Antimicrobial agents and chemotherapy **57**, 1238 (2013).

[52] Y. Zhang, Y. Zhu, A. Gupta, Y. Huang, C. K. Murray, M. S. Vrahas, M. E. Sherwood, D. G. Baer, M. R. Hamblin, and T. Dai, *Antimicrobial blue light therapy for multidrug-resistant acinetobacter baumannii infection in a mouse burn model: implications for prophylaxis and treatment of combat-related wound infections,* The Journal of infectious diseases **209**, 1963 (2014).

[53] K. Kang-Hsin Wang, S. Mitra, and T. H. Foster, *A comprehensive mathematical model of microscopic dose deposition in photodynamic therapy,* Med. Phys. **34**, 282 (2007).

[54] M. M. Kim, A. A. Ghogare, A. Greer, and T. C. Zhu, *On the in-vivo photochemical rate parameters for pdt reactive oxygen species modeling,* Phys. Med. Biol. **62**, R1 (2017).

[55] T. C. Zhu, J. C. Finlay, X. Zhou, and J. Li, *Macroscopic modeling of the singlet oxygen production during pdt,* in *Optical Methods for Tumor Treatment and Detection: Mechanisms and Techniques in Photodynamic Therapy XVI*, Vol. 6427 (International Society for Optics and Photonics, 2007) p. 642708.

[56] L. Ljung, *System Identification - Theory for the User* (Prentice-Hall, 1987).

[57] M. R. Amin and R. T. Faghih, *Sparse deconvolution of electrodermal activity via continuous-time system identification,* IEEE. Trans. Biomed. Eng. **66**, 2585 (2019).

[58] A. Patel, J. Li, B. Finegan, and M. McMurtry, *Aortic pressure estimation using blind identification approach on single input multiple output nonlinear wiener systems,* IEEE. Trans. Biomed. Eng. **65**, 1193 (2018).

[59] Y. Gu, Y. Yang, J. Dewald, F. C. T. Van der Helm, A. C. Schouten, and H. Wei, *Nonlinear modeling of cortical responses to mechanical wrist perturbations using the NARMAX method,* IEEE. Trans. Biomed. Eng. **In Press** (2020).

[60] K. K.-H. Wang, J. C. Finlay, T. M. Busch, S. M. Hahn, and T. C. Zhu, *Explicit dosimetry for photodynamic therapy: macroscopic singlet oxygen modeling,* J. Biophotonics **3**, 304 (2010).

2

BLUE LIGHT THERAPY TO TREAT CANDIDA INFECTION WITH COMPARISONS OF THREE WAVELENGTHS: AN IN VITRO STUDY

*Anti-fungal blue light (ABL) therapies have been widely studied to treat various microbial infections in the literature. The blue light with wavelengths ranging from 400-nm to 470-nm has been reported to be effective to inhibit various kinds of bacteria and fungi. The existing studies usually report the viability rates of the pathogens under the irradiation of the blue light with different dosage parameters. However, to the best of our knowledge, there is still no work especially focusing on studying the effect of ABL therapies on treating candidal vaginitis, where it is important to study the viability of both the candida albicans (*C. albicans*) and the human vaginal epithelial cells. It is the purpose of this work to conduct ABL experiments on both of these two cells, analyse the effects and determine the best ABL wavelength out of three candidates, i.e. 405-nm wavelength, 415-nm wavelength and 450-nm wavelength. The viability rates of the* C. albicans *and the human vaginal epithelial cells irradiated by the three blue LED light sources were measured, whose irradiance (power density) were all set to 50 mW/cm^2. The dynamic viability models of the* C. albicans *and the epithelial cells were also built based on the experimental data. Moreover in this work, we also built a functional relationship between the viability of these two types of cells, by which we further compared the effects of the blue light irradiation on both the* C. albicans *and vaginal epithelial cells. The experimental data showed that when an approximately 80% inhibiting rate of the* C. albicans *was achieved, the survival rates of the epithelial cells were 0.6700, 0.7748 and 0.6027, respectively for the case of 405-, 415- and 450-nm irradiation. On the other hand, by simulating the functional relationship between the viability of the two types of cells, the survival rates of the epithelial cells became 0.5783, 0.6898 and 0.1918 respectively using 405-, 415- and 450-nm, when the* C. albicans

was completely inhibited. Therefore, both the experimental data and the model simulation results have demonstrated that the 415-nm light has a more effective anti-fungal result with less damage to the epithelial cells than the 405- and 450-nm light.

2

2.1. INTRODUCTION

The *C. albicans* is widely found in nature and commonly occurs as a superficial infection on mucous membranes in mouth or vagina [1, 2], which is the most common pathogen of fungal diseases of human [3]. What's more, it can also cause visceral or systemic infections [4, 5].

As a common disease by fungal infection, candidal vaginitis is a common cause of gynecologic infections in United States, with an increased incidence seen in pregnant women, diabetic women, and women receiving antibiotic or corticosteroid therapy [6, 7]. By the age of 25 half of all college women have at least one physician-diagnosed episode of vulvovaginal candidiasis. It is less common in post-menopausal women. In other populations, at least one episode of vulvovaginal candidiasis is reported in up to 75 percent of premenopausal women [8]. About 90% of these infections are caused by *C. albicans* [9].

At present, antibiotics such as penicillins, cephalosporins and anti-fungal drugs and oral azole agents are the main treatments for fungal infections [8]. However, *C. albicans* is showing an increasing resistance for these drugs [10, 11]. Therefore it is crucial to find new ways to treat fungal infections.

As a non-antibiotic approach, photobiomodulation therapy (PBMT) has been investigated as an alternative method for localized infectious diseases. The advantages of PBMT include rapid action and avoidance of drug resistance [12, 13]. The anti-microbial effects of PBMT are highly depend on the light wavelength. Red light has no anti-microbial effects; while the green one has minor of that [14]. In contrast, researches have shown that some specific wavelengths of blue light and ultraviolet light have obvious anti-microbial effects on most pathogenic microbes [15, 16]. The use of UV light has limitations due to its detrimental effects on host cells [17] and can cause mutagenic oxidative DNA damage [18]. Therefore, blue light (400-nm to 470-nm) is more suitable than other wavelengths in treating microbial infections.

At present, lasers and LEDs are two main types of ABL light sources. Compared with lasers, LEDs, especially in visible light range, are much more affordable and safer than lasers, and can be easily integrated into an array to treat much larger surfaces [19]. They are hence appealing alternatives for lasers in various light therapies [19–21], including ABL therapies [16, 22]. For instance, in [23], it was reported that the 405-nm wavelength LED light was able to entirely eliminate two species of candida strains after 30-minute irradiation with the irradiance of 280 mW/cm^2. In another work [22], the antimicrobial effect of 415-nm LED on *C. albicans* was studied in vitro and in vivo. In [16], it was proved that 405-nm ABL has an obvious inactivating effect on *C. albicans* and other microbes. Another research further compared *C. albicans* with human keratinocytes cells, which indicates that the inactivation rate of *C. albicans* by the 415-nm wavelength light is 42-fold higher than that of the human cells [22]. Different wavelengths of blue LED light were also studied and compared in [24]. It demonstrated that the bactericidal effect of 405-nm wavelength LED light source was better than that of 470-nm wavelength.

However, the inhibiting effects of three mostly used wavelengths, i.e. 405-, 415- and 450-nm, have not been compared yet. Moreover, for any ABL therapies, besides studying anti-fungal or bactericidal effects, it is also very important to study the potential damage of the therapies to the human host cells. Especially for treating candidal vaginitis as studied in this work, it is essential to study the inactivating effects of both the *C. albicans* and vaginal epithelial cells. However, in the existing studies, the inhibiting effect of blue light on these human host cells has not been reported yet. On the other hand, the quantitative models that describe the fungal viability changes over time have not yet attracted much research attentions. Although there are some existing mathematical dynamic models of microbial inactivation, the functional relationship between the viability of microbial pathogens and human host cells have not yet been investigated. The importance of such a function is that it can further facilitate quantitatively comparing the effects of the blue light irradiation on these cells.

The contributions of the current work are three-folds. First, we conducted blue light stimulation experiments on both *C. albicans* and human vaginal epithelial cells, using three different wavelengths, including 405-nm, 415-nm and 450-nm. Second, we built dynamic viability models of the *C. albicans* and epithelial cells, by which we further analyzed the effects of the light stimulation to both cells. Third, the optimal ABL wavelength for treating candidal vaginitis has been determined via both experiments and mathematical modeling.

2.2. MATERIALS AND METHODS

2.2.1. FUNGAL AND CELL STRAINS

The *C. albicans* used in this work is the 3147 (IFO 1594) strains (ATCC,US). The fungal strains was cultured in tryptic soy broth (TSB) at 26^oC. The human vagina cell line VK2/E6E7 (ATCC, USA) was incubated in Dulbecco's modified Eagle's medium (Gibco, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum at 37^oC in a humidied atmosphere with 5% CO_2.

2.2.2. ANTI-FUNGAL ASSAYS

The viability of the fungi was estimated by colony counting in terms of colony forming units (CFU). After the concentration of fungi reached 10^7 yeast cells/ml. The fungal suspension was diluted by 10^4-fold with sterilized water and then spread on the tryptic soy agar (TSA). The plates were divided into control groups and treatment groups. The treatment groups were irradiated at $50mw/cm^2$ for various time, while the control groups were kept in the dark. Both the treatment groups and the control groups were cultured over 24-48 hours at 26^oC before the colony counting. Triplicate experiments were performed at each sampling time for both control and treatment groups. The raw data were processed to produce the mean and standard deviation of the viability rate at each sampling time. The significance of the viabllity of the fungi was tested by the Student's t-test.

Table 2.1: The LEDs used in this wirk and their main parameters, where FWHM represents full width at half maximum.

type	peak wavelength(nm)	FWHM (nm)
Kingbright ATDS3534UV405B	405	15
LUMILEDS LUXEON LHUV-0415-A070	415	13.7
Cree XLampXPE2	450	20

2.2.3. CELL VIABILITY ASSAYS

The viability of the epithelial cells were evaluated using a CCK-8 tool kit (Beyotime Institute of Biotechnology, China), which is based on the cleavage of the tetrazolium salt WST-8 by mitochondrial dehydrogenase in viable cells. The cells were seeded in 9 wells of 96-well plate in the shape of 3 by 3 as the treatment group. Another 9 wells in the same 96-well plate were seeded by the cells as the control group. After culturing overnight, the concentration is about 10000 cells/well. Then the treatment group was irradiated by blue light for various time, while the control group was covered in the dark. Then the cell counting kit-8 (CCK-8) was added to the culture medium with a ratio of 1:10. After a one-hour incubation at 37^oC, the level of absorbance of 450-nm light was measured by Bio-tek Synergy™ HT microplate reader (Bio-tek, US), in terms of the optical density (OD). The viability rate N_c was evaluated as follows.

$$N_v = OD_{treatment}/OD_{control} \times 100\%, \tag{2.1}$$

where the $OD_{treatment}$ and $OD_{control}$ stands for the optical density of the treatment group and the control group, respectively. Triplicate experiments were performed at each sampling time. The raw data were processed to produce the mean and standard deviation of the viability rate at each sampling time. The significance of the viabllity of the cells was tested by the Student's t-test.

2.2.4. LED LIGHT SOURCES

In this work, we used three types of blue LEDs as listed in Tab.2.1. Both the 405-nm light source and the 415-nm light source were made of a 4 by 4 LED array, arranged in a 1.5 by 1.5 cm square. The 450-nm light source was made of a 3 by 3 LED array arranged in a 1 by 1 cm square. The spectral power density (SPD) of the three blue light sources was measured by a Maya2000Pro spectrometer (Ocean Optics, US), which is shown in Fig. 2.1. The LEDs always contains some UVA content. More specifically, the 405-nm LED contains 36.31% UVA in terms of its power in this band over that of the entire spectral range; the amount of UVA in the 415- and 450-nm LED are 3.73% and 1.13% respectively.

The light power density and uniformity was measured using a PM100D power/energy meter with a S120VC probe (Thorlabs, US) in the experiments. In a round light spot with a 35mm diameter, the difference of light density between center and edge is less than 10%. Fungi were cultured on the medium agar so the light power density can be measured directly. The 50 mW/cm^2 optical power density were calibrated by adjusting the height of the light source or the electrical power of the LED. Since the cells were adherent cultured in a 96-well plate below the liquid medium (DMEM), to evaluate the light power

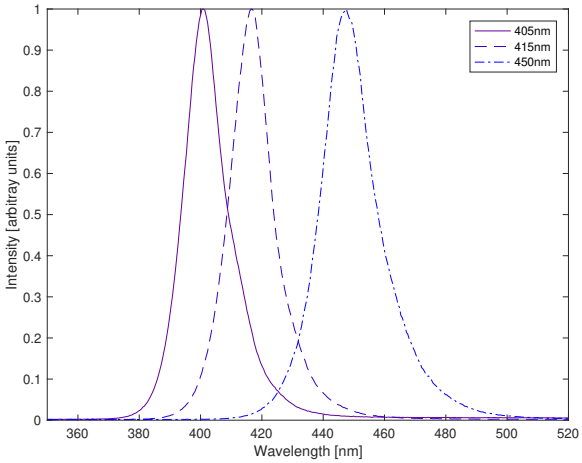


Figure 2.1: The measuring results of the actual wavelengths of 405-nm, 415-nm and 450-nm.

Table 2.2: Estimated effective light irradiance on cells.

Irriadiance (mW/cm^2)	P1	P2	P3	P1-(P2-P3)
405-nm	60.2	54.8	44.3	49.7
415-nm	67.6	62.5	45.0	50.1
450-nm	70.6	63.5	43.2	50.3

density on the cell surface, it is necessary to measure the absorption of the medium. The methods were illustrated as follows.

The attenuation of light while travelling trough a material follows the Lambert-Beer law [25]

$$I(x) = I(0)e^{-kx}, \tag{2.2}$$

where $I(0)$ is the irradiance at the sample surface; $I(x)$ is the irradiance at the sample depth x; and k is the attenuation coefficient depends on the material. However, the culture medium was a mixture solution, the attenuation coefficient cannot be estimated experimentally. So we measured the irradiances at the sample surface and minus it by the absorption of the culturing medium.

Fig. 2.2 gives a sketch of the measuring approach. The measuring results were listed in Tab. 2.2. The light density on the cells surface was estimated by

$$P = P_1 - (P_2 - P_3) \tag{2.3}$$

Where P_1, P_2 and P_3 were measured under three different conditions as illustrated before. $P_2 - P_3$ measures the light absorption of the medium, so $P = P1 - (P2 - P3)$ was the light density on the cell surface, above the dish bottom while below the culture medium.

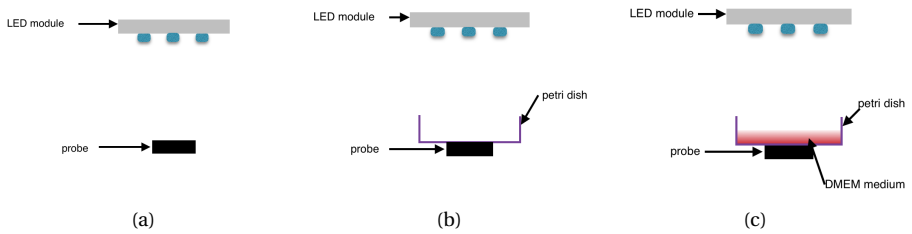


Figure 2.2: Three different conditions of measuring the irradiance. P_1 is the measured directly with the probe. P_2 is measured below a 35mm patri dish with lid opened. P_3 is measured below a 35mm patri dish filled with 2 mm thickness culturing medium without the lid.

2.2.5. FITTING DYNAMIC MODELS FROM THE VIABILITY DATA

In thermal and non-thermal inactivation of vegetative microorganism, there are four common types of the survival curves: linear curves, curves with a shoulder, curves with a tailing and sigmoidal curves [26]. In this study, there is no significant tailing phase in the survival curves. Therefore we fitted the curves of the survival rates with two biology models. Model 1 takes the form as first order kinetics [27]

$$N(t) = N(0)e^{-kt}, \tag{2.4}$$

where k is the decaying rate coefficient which can be simply solved by least-squares method. Here, $N(t)$ is monotonic decline with t increasing. However, the main mechanism of ABL is the reactive oxygen species (ROS) accumulating. The ROS are reactive chemical species containing oxygen, which may result in damage to cell structures. Cells become more and more susceptible while the ROS accumulating caused by light. Therefore, a shoulder phase representing the ROS accumulating is necessary in the model fitting. Here, model 2 is a piecewise function [28], which indicates a shoulder of the curve, and takes the following form

$$N(t) = \begin{cases} N(0), & t \le \tau \\ N(0)e^{-k(t-\tau)}, & t > \tau \end{cases} \tag{2.5}$$

where $N(t)$ is the survival rate at time t; k is the decaying rate coefficient and τ is the time constant of when the inactivation starts. These two parameters can be solved as follows.

First, fix τ and solve k by least-squares method. The model can be rewritten as a least-squares problem

$$\boldsymbol{y} = -k\boldsymbol{x} \tag{2.6}$$

where $\boldsymbol{y} = [y_1, y_2 \cdots y_T]$, $\boldsymbol{x} = [x_1, x_2 \cdots x_T]$, with the following definition.

$$\begin{cases} y_t = \ln[N(t)/N(0)] \\ x_t = t - \tau \end{cases} \tag{2.7}$$

while $t > \tau$. The decaying rate coefficient k can be simply solved by least-squares method, then select the minimal MSE, its corresponding τ and k to be the final solution from all candidates. This method is illustrated in Fig. 2.3.

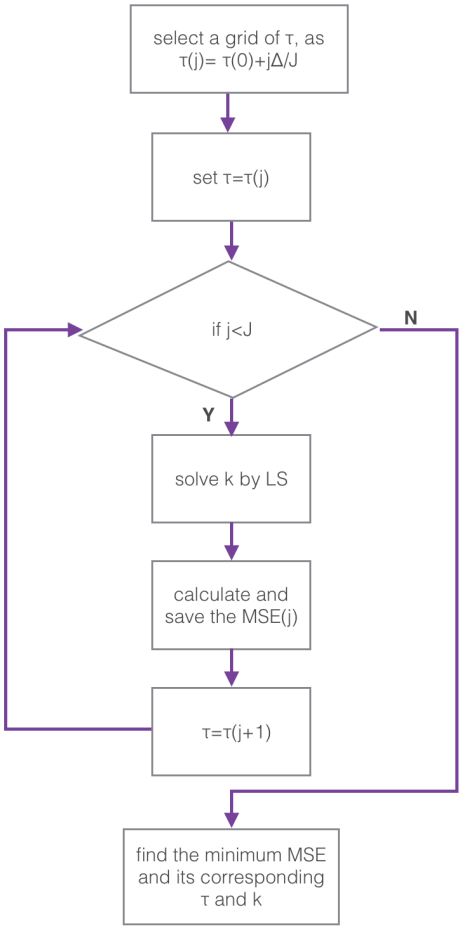


Figure 2.3: The parameters estimation of model 2. The index of the candidates is denoted as j, where $j = 1,2,\cdots,J$; Δ is the target range and $\tau(0)$ is the starting point.

2.2.6. FUNCTIONAL RELATIONSHIP BETWEEN THE VIABILITY OF THE *C. albicans* AND VAGINAL EPITHELIAL CELLS

To analyse the effects of ABL therapy on both the *C. albicans* and vaginal epithelial cells, it is necessary to establish the mathematical relationship between the viability of these two types of cells. Based on model 2, one can derive that

$$N_c(t) = N_c(0)e^{-k(t-\tau_c)},\ t>\tau_c \tag{2.8}$$
$$N_v(t) = N_v(0)e^{-k(t-\tau_v)},\ t>\tau_v, \tag{2.9}$$

where the subscript means the type of cells, i.e. *C. albicans* or vaginal epithelial cells. In the simulation, $N(t)$ is the survival rate, therefore $N_c(0), N_v(0) = 1$ and Eq. (2.8)-(2.9) can be rewritten as

$$N_c = \begin{cases} 1, & t \le \tau_c \\ e^{-k_c(t-\tau_c)}, & t > \tau_c \end{cases} \tag{2.10}$$

$$N_v = \begin{cases} 1, & t \le \tau_v \\ e^{-k_v(t-\tau_v)}, & t > \tau_v \end{cases} \tag{2.11}$$

Note that in Eq. (2.10)-(2.11), the constants τ_c, τ_v determine the time when the viability of the cells starts to decrease. The difference between τ_c and τ_v indicates the inactivation of which type of cell starts first.

We first consider the case of $\tau_c \le \tau_v$. For $t < \tau_c$, it is clear that

$$N_v, N_c = 1 \tag{2.12}$$

For $\tau_c \le t \le \tau_v$,

$$N_c = N_c e^{-k(t-\tau_c)} \tag{2.13}$$

$$N_v = 1 \tag{2.14}$$

For $\tau_c \le \tau_v < t$, solve t from Eq. (2.10)

$$t = -\frac{\ln N_c}{k_c} + \tau_c \tag{2.15}$$

Plug t into Eq. (2.11)

$$\begin{aligned} N_v &= e^{-k_v(t-\tau_v)} \\ &= e^{-k_v(-\frac{\ln N_c}{k_c}+\tau_c-\tau_v)} \\ &= e^{-k_v(\tau_c-\tau_v)} N_c^{\frac{k_v}{k_c}} \end{aligned} \tag{2.16}$$

Finally, from Eq. (2.12)-(16)

$$N_v = \begin{cases} 1, & N_c \ge e^{-k_c(\tau_v-\tau_c)} \\ e^{-k_v(\tau_c-\tau_v)} N_c^{\frac{k_v}{k_c}}, & N_c < e^{-k_c(\tau_v-\tau_c)} \end{cases} \tag{2.17}$$

On the other hand, for the case of $\tau_v < \tau_c$ can be derived in the same way as follows. For $t < \tau_v$,

$$N_v, N_c = 1$$

For $\tau_v \le t \le \tau_c$,

$$N_v = e^{-k_v(t-\tau_v)}$$

For $\tau_v < \tau_c < t$, N_v takes the form as Eq. (16). Thus

$$N_v \begin{cases} \in \left[e^{-k_v(\tau_c-\tau_v)}, 1\right], & N_c = 1 \\ = e^{-k_v(\tau_c-\tau_v)} N_c^{\frac{k_v}{k_c}}, & N_c < 1 \end{cases} \tag{2.18}$$

Here, the functional relationship between the viability of two types of cells is derived as aforementioned, for all the six assumptions. This functional relationship offers the opportunity to estimate the treatment efficacy and host cell damage, which in turn helps to optimize the light dosage.

From Eq. (2.17) and Eq. (2.18), the derivative of N_v which respect to N_c when $N_v < 1$ is

$$\frac{dN_v}{dN_c} = e^{-k_v(\tau_c - \tau_v)} \frac{k_v}{k_c} N_c^{\frac{k_v}{k_c}-1}. \tag{2.19}$$

If this gradient above satisfy

$$\frac{k_v}{k_c} < 1, \tag{2.20}$$

then, $\lim_{N_c \to 0} \frac{dN_v}{dN_c} = \infty$. This function indicates that N_v can be much lager than zero as N_c goes to zero. Eqs. (2.19-2.20) give a mathematical condition which leads to a acceptable treatment outcome, i.g., if this condition is not satisfied, the light treatment should not be applied.

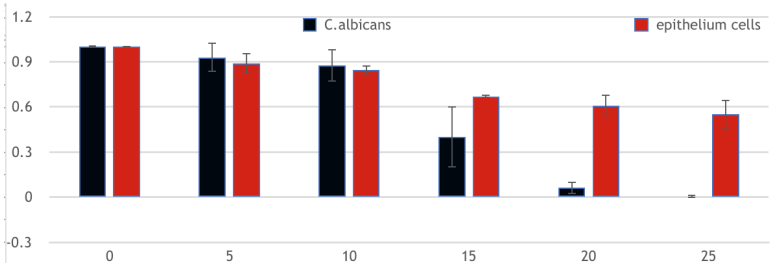
2.3. RESULTS

2.3.1. THE ABL INACTIVATION OF THE *C. albicans*

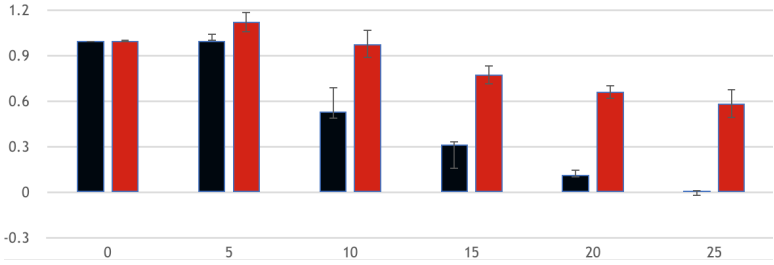
The viability of the *C. albicans* showed an exponential relation between the viability rate and the irradiated time Fig. 2.4. After a 50 mW/cm^2 optical power of density was delivered, no significant growth inhibition of the *C. albicans* was observed within 10 minutes, 5 minutes and 45 minutes respectively exposed by 405-, 415- and 450-nm wavelength ($P > 0.05$). The significant growth inhibition was observed in other higher doses ($P < 0.05$): irradiation time over 10 minutes, 5 minutes and 45 minutes respectively using 405- , 415- and 450-nm wavelength. This phenomenon indicates a ROS accumulating process. Once the concentration of ROS reached a specific level, the inactivating effects will start. In the end of the experiments, the survival rates were 0.0056 and 0.0089 respectively at the exposure of 405- and 415-nm wavelength, which can be considered as the total apoptosis with considering the CFU counting errors. In the end of the 450-nm wavelength experiments, 0.2344 of the *C. albicans* were survived, which indicates a less anti-fungal performance.

2.3.2. THE ABL EFFECTS OF THE VAGINAL EPITHELIAL CELLS

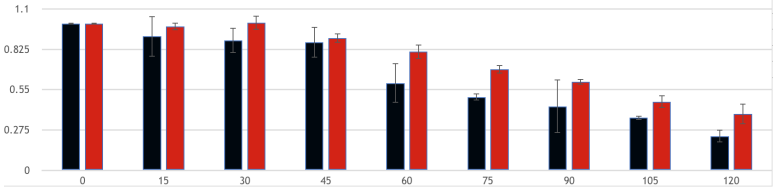
The ABL effects of the vaginal epithelial cells were similar to that of the *C. albicans*. With the same fluence of 50 mW/cm^2, no significant growth inhibition of the epithelial cells was observed before 5 minutes, 10 minutes and 45 minutes respectively exposed by 405- , 415- and 450-nm wavelength ($P > 0.05$). The significant growth inhibition was observed in other higher doses($P < 0.05$): irradiation time exceed 5 minutes, 10 minutes and 45 minutes respectively using 405- , 415- and 450-nm wavelength. In the end of the experiments, the survival rates were 0.5700, 0.5861 and 0.3853 respectively irradiated by 405- , 415- and 450-nm wavelength.



(a)



(b)



(c)

Figure 2.4: The survival rates for the *C. albicans* and vaginal epithelial cells in the irradiation of three wavelengths. (a) 405-nm. (b) 415-nm. (c) 450-nm.

2.3.3. FITTING DYNAMIC MODEL

It is clear to show the advantage of the irradiation of 415-nm over that of 450- and 405-nm in Fig. 2.5. The abscissa axis is reversed for better showing inactivation process of the decreasing viability. While the survival rates of the *C. albicans* irradiated by different wavelengths were same, the survival rates of the epithelial cells were higher at the irradiation of 415-nm than that of the other two wavelengths. The slope for a liner fitting is 0.453 for 405-nm wavelength, 0.356 for 415-nm and 0.726 for 450-nm wavelength respectively. The slope reflects the dying speed of the epithelial cells by the same inactivation rate of *C. albicans*. Hence 415-nm is the best wavelength out of the three.

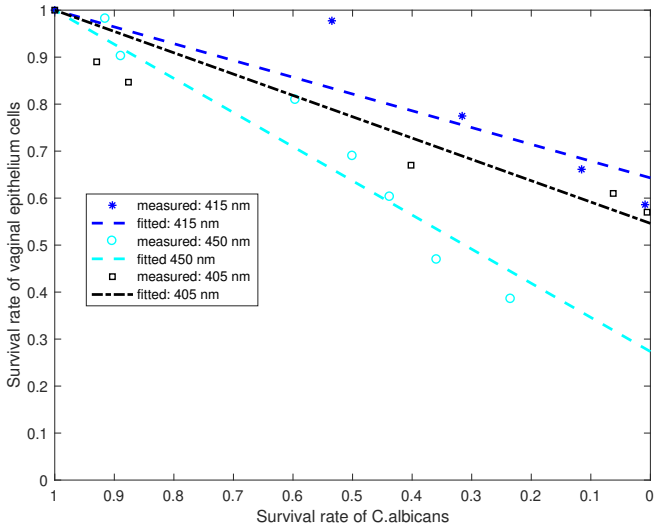


Figure 2.5: Comparison of survival rates and its decreasing trends of linear function.

The simulation results by model 2 are shown in Fig. 2.6, the MSE errors compared with model 1 are listed in Tab. 2.4 and the fitting parameters are listed in Tab. 2.3. In all cases, model 2 is more accurate than model 1 in terms of the lower MSE error. The better fitting accuracy of model 2 also indicates that the survival curves has a shoulder phase.

Table 2.3: The estimated parameters of model 2, where C. A denotes the *C. albicans* and V. E. denotes the vaginal epithelial cells.

	405-nm		415-nm		450-nm	
	τ	k	τ	k	τ	k
C. A.	769	6.93e-3	450	3.80e-3	2017	2.64e-4
V. E.	56	4.12e-4	526	5.81e-4	2307	1.85e-4

Table 2.4: The fitting MSE by the two models, where C. A denotes the *C. albicans* and V. E. denotes the vaginal epithelial cells.

	405-nm		415-nm		450-nm	
	model 1	model 2	model 1	model 2	model 1	model 2
C.A.	0.0412	0.0034	0.0220	0.0038	0.0078	0.0031
V.E.	7.7304e-4	7.1920e-4	0.0111	0.0028	0.0087	4.5370e-04

2.3.4. FUNCTIONAL RELATIONSHIP BETWEEN C.*abicans* AND HUMAN VAGINAL EPITHELIAL CELLS BASED ON THE EXPERIMENTAL DATA

Based on the parameters in Tab. 2.3, the factional relationship of the two types of cells defined by Eq. (2.17) and Eq. (2.18) and its decreasing trends are shown in Fig. 2.7. The slopes of the liner functions are 0.4217, 0.3102 and 0.8082 for 405- , 415- and 450-nm, respectively. Which indicates that 415-nm is the best wavelength out of these three candidates. Compare to the previous fitted linear model by the experimental data in Fig. 2.5, the slopes are similar with a difference less than 12%.

2.4. DISCUSSION

With a short time exposure, neither cells or fungi were inactivated as shown in Fig. 2.4, which means there is a 'shoulder' phase existed. The similar phenomenon was also reported in other studies. In the exposure of 405-nm anti-bacterial blue light, during the first 10 minutes, no obvious apoptosis occurred in all bacteria [15]. This can be attributed to mechanism of ABL therapies. Though it is still not fully understood, a common acceptable hypothesis is that ABL excites naturally existing endogenous photosensitizers and leads to the generation of cytotoxic reactive oxygen species (ROS) [22]. When the ROS accumulates to a lethal concentration, the apoptotic pathway is activated and hence the inactivation can be observed. This explains why the inactivation always starts after the light delivering. Researches have determined some types of the endogenous photosensitizers such as uroporphyrin and coproporphyrin in some common bacteria [29–31]. However. To the best of our knowledge, the endogenous photosensitizers in fungi have not been determined yet. The amount and types of the endogenous photosensitizers can be different from that of bacteria.

It is observed in the experiment that a small dose of blue light can even stimulate the growth of the cells. Within five minutes irradiation of 415-nm blue light, the viability of cells was increased by 10%. This phenomenon could be attributed to the cellular stress response. The cellular stress response is a wide range of molecular changes that cells undergo in response to environmental stressors. There were few researches about stress response to blue light irradiation, but in an UVA light exposure experiment [32], the exposure of UVA induced proteins comprised other several stress-related proteins and further stimulated the growth of the basidiomycete Nidula niveotomentosa. The similar response may also happen in the exposure of blue lights whose wavelength is close to UVA.

The *C. albicans* and vaginal epithelial cells are more susceptible to 405- , 415- and 450-nm wavelength blue light, which is similar to the research that shorter wavelength

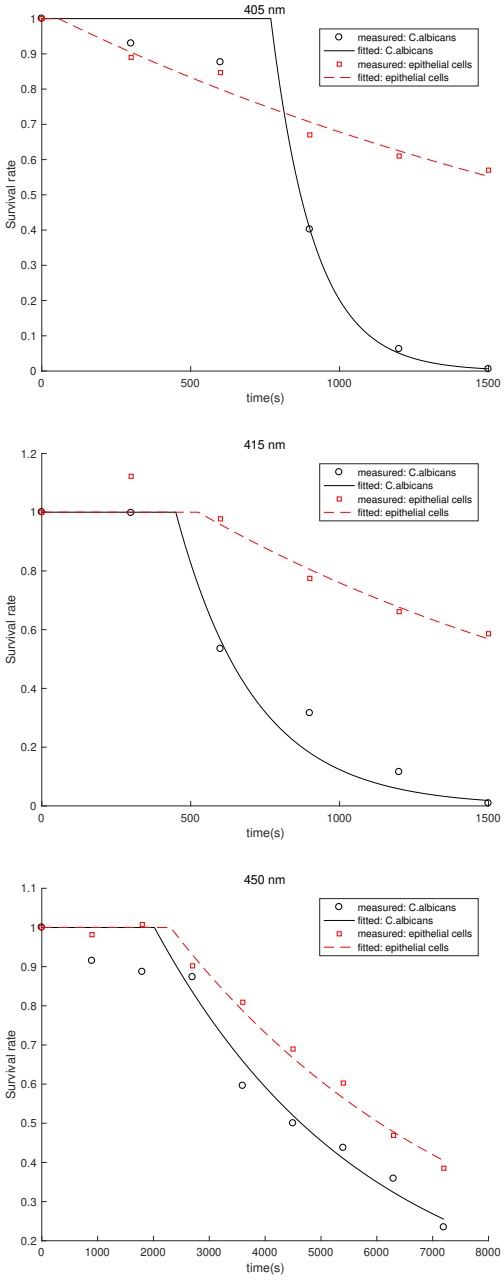
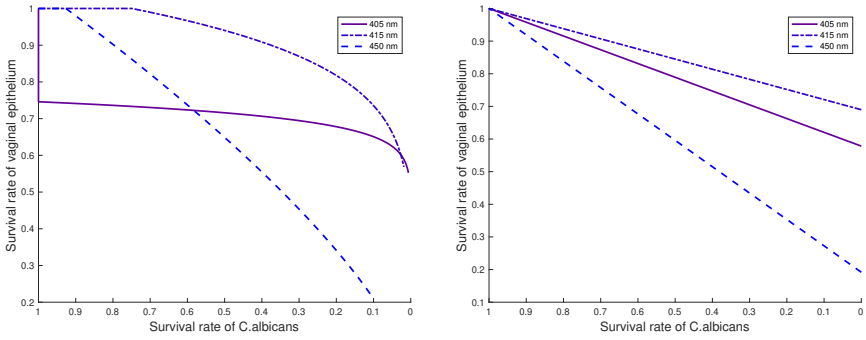


Figure 2.6: Simulation results of model 2. (a) 405-nm. (b) 415-nm. (c) 450-nm.



(a) (b)

Figure 2.7: (a) Functional relationship of the survival rates by taking the viability of the *C. albicans* as variable. (b) The decreasing trends of a linear function.

(400-nm to 450-nm) irradiation of blue light is more harmful to retina [33]. It is also well known that UV is an ionized irradiation and is hence more harmful to most organisms. Therefore, it becomes a trade-off on ABL therapy in this study; i.e. if seeking a higher inactivating speed of *C. albicans* by a shorter wavelength, the epithelial cells are also inactivated in a higher rate. So there should be an optimal ABL wavelength which can inactivate *C. albicans* in a proper speed and is less harmful to the epithelial cells. In this study, ABL of 415-nm wavelength was proofed better than the other two wavelengths.

In the fitted dynamic model, in all cases, the value $\frac{k_v}{k_c}$ is less than 1 which satisfies the Eq. (2.20), which in turn leads to the derivative $\frac{dN_v}{dN_c} = \infty$. However, in the 450-nm experiments N_v is close to 0 in the ending phase while in other two experiments N_v is over 0.5 in the end. This can be attributed to the value of $\frac{k_v}{k_c}$. In both 405-nm and 415-nm experiments, the value of $\frac{k_v}{k_c}$ are less than 0.17. However in 450-nm experiments, this value is increased to 0.70. This value indicates the sensitivity of N_v in terms of the change of N_c. Thus, the lower value of $\frac{k_v}{k_c}$ can lead to a larger N_v in the end.

On the other hand, with the estimation of the parameter τ in model 2, it is clear that in the 415-nm experiment, the apoptosis of the *C. albicans* starts earlier than the vaginal epithelial cells. In contrast, the apoptosis of the epithelial cells starts earlier in the 405-nm experiments. This can be attributed to the DNA damage caused by UV [18]. From the measuring results, the 405-nm light source contains 36.31% UVA. Here, we can make an hypothesis that this portion of UVA irradiation cause the DNA damage of the epithelial cells, thus the the apoptosis of the epithelial cells starts earlier than the *C. albicans*. As in 450-nm experiments, there is not an obvious time difference of the apoptosis start. Anyway, the 415-nm blue light show its advantage both in terms of its smallest slope in the survival curves and the apoptosis of the *C. albicans* starts earlier than the epithelial cells.

2.5. CONCLUSION

In this work, the inactivating effects of the *C. albicans* and vaginal epithelial cells by the irradiation of 405- , 415- and 450-nm were studied. Moreover, the dynamic viability models and functional relationship of these two types of cells were built based on the experimental data. The results indicate that, in vitro, the 415-nm light source has a more effective anti-fungal function with less damage to the human host cells than 450-nm and 405-nm wavelength. The optimal wavelength and fluence found in this study can be helpful to design the therapeutic devices and the doses used in treating candidal vaginitis. A potential extension of our current work is to further investigate the optimal dosage of treating this disease in clinical experiments, and to model the dynamic changes of the viability of the fungi as well as that of the human epithelial cells.

REFERENCES

[1] T. Arendorf and D. Walker, *The prevalence and intra-oral distribution of candida albicans in man,* Arch. Oral Biol. **25**, 1 (1980).

[2] R. A. Calderone and W. A. Fonzi, *Virulence factors of candida albicans,* Trends Microbiol. **9**, 327 (2001).

[3] L. J. Douglas, *Candida biofilms and their role in infection,* Trends Microbiol. **11**, 30 (2003).

[4] A. Erdogan and S. S. C. Rao, *Small intestinal fungal overgrowth,* Curr. Gastroenterol. Rep. **17**, 16 (2015).

[5] R. Singh and A. Chakrabarti, *Invasive candidiasis in the southeast-asian region,* in *Candida albicans: Cellular and Molecular Biology* (Springer, 2017) pp. 25–40.

[6] A. O. Berg, F. E. Heidrich, S. D. Fihn, J. J. Bergman, R. W. Wood, W. E. Stamm, and K. K. Holmes, *Establishing the cause of genitourinary symptoms in women in a family practice: comparison of clinical examination and comprehensive microbiology,* Jama **251**, 620 (1984).

[7] C. Morris and D. F. Morris, *Normal'vaginal microbiology of women of childbearing age in relation to the use of oral contraceptives and vaginal tampons,* J. Clin. Pathol. **20**, 636 (1967).

[8] J. D. Sobel, *Vaginitis,* N. Engl. J. Med. **337**, 1896 (1997).

[9] F. C. Odds *et al., Candida and candidosis: a review and bibliography* (Bailliere Tindall, 1988).

[10] M. A. Pfaller, *Antifungal drug resistance: mechanisms, epidemiology, and consequences for treatment,* Am. J. Med. **125**, S3 (2012).

[11] P. Vandeputte, S. Ferrari, and A. T. Coste, *Antifungal resistance and new strategies to control fungal infections,* Int. J. Microbiol. **2012** (2011).

[12] Z. A. Kanafani and J. R. Perfect, *Resistance to antifungal agents: mechanisms and clinical impact,* Clin. Infect. Dis. **46**, 120 (2008).

[13] T. Maisch, *A new strategy to destroy antibiotic resistant microorganisms: antimicrobial photodynamic treatment,* Mini-Rev. Med. Chem. **9**, 974 (2009).

[14] S. Kim, J. Kim, W. Lim, S. Jeon, O. Kim, J.-T. Koh, C.-S. Kim, H. Choi, and O. Kim, *In vitro bactericidal effects of 625, 525, and 425 nm wavelength (red, green, and blue) light-emitting diode irradiation,* Photomed. Laser Surg. **31**, 554 (2013).

[15] M. Maclean, S. J. MacGregor, J. G. Anderson, and G. Woolsey, *Inactivation of bacterial pathogens following exposure to light from a 405-nanometer light-emitting diode array,* Appl. Environ. Microbiol. **75**, 1932 (2009).

[16] W. J. Trzaska, H. E. Wrigley, J. E. Thwaite, and R. C. May, *Species-specific antifungal activity of blue light,* Sci. Rep. **7**, 4605 (2017).

[17] T. Dai, M. S. Vrahas, C. K. Murray, and M. R. Hamblin, *Ultraviolet c irradiation: an alternative antimicrobial approach to localized infections?* Expert Rev. Anti-Infect. Ther. **10**, 185 (2012).

[18] P. O'donovan, C. M. Perrett, X. Zhang, B. Montaner, Y.-Z. Xu, C. A. Harwood, J. M. McGregor, S. L. Walker, F. Hanaoka, and P. Karran, *Azathioprine and uva light generate mutagenic oxidative DNA damage,* Science **309**, 1871 (2005).

[19] J. Dong and D. Xiong, *Applications of light emitting diodes in health care,* Ann. Biomed. Eng. **45**, 2509 (2017).

[20] D. Barolet, *Light-emitting diodes (LEDs) in dermatology,* Semin. Cutan. Med. Surg., **27**, 227 (2008).

[21] I. Figueira, A. C. Sousa, A. Machado, F. Habib, L. Soares, and A. Pinheiro, *Clinical study on the efficacy of LED phototherapy for pain control in an orthodontic procedure,* Lasers Med. Sci. **34**, 479 (2019).

[22] Y. Zhang, Y. Zhu, J. Chen, Y. Wang, M. E. Sherwood, C. K. Murray, M. S. Vrahas, D. C. Hooper, M. R. Hamblin, and T. Dai, *Antimicrobial blue light inactivation of candida albicans: in vitro and in vivo studies,* Virulence **7**, 536 (2016).

[23] C. Tsutsumi-Arai, Y. Arai, C. Terada-Ito, Y. Takebe, S. Ide, H. Umeki, S. Tatehara, R. Tokuyama-Toda, N. Wakabayashi, and K. Satomura, *Effectiveness of 405-nm blue LED light for degradation of candida biofilms formed on PMMA denture base resin,* Lasers Med. Sci. , 1 (2019).

[24] J. S. Guffey and J. Wilborn, *In vitro bactericidal effects of 405-nm and 470-nm blue light,* Photomed. laser surg. **24**, 684 (2006).

[25] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light* (Elsevier, 2013).

[26] R. Xiong, G. Xie, A. Edmondson, and M. Sheard, *A mathematical model for bacterial inactivation,* Int. J. Food Microbiol. **46**, 45 (1999).

[27] H. Chick, *An investigation of the laws of disinfection,* Epidemiol. Infect. **8**, 92 (1908).

[28] R. Buchanan, M. Golden, and R. Whiting, *Differentiation of the effects of ph and lactic or acetic acid concentration on the kinetics of listeria monocytogenes inactivation,* J. Food Prot. **56**, 474 (1993).

[29] T. Dai, A. Gupta, Y.-Y. Huang, R. Yin, C. K. Murray, M. S. Vrahas, M. E. Sherwood, G. P. Tegos, and M. R. Hamblin, *Blue light rescues mice from potentially fatal pseudomonas aeruginosa burn infection: efficacy, safety, and mechanism of action,* Antimicrob. Agents Chemother. **57**, 1238 (2013).

[30] M. R. Hamblin, J. Viveiros, C. Yang, A. Ahmadi, R. A. Ganz, and M. J. Tolkoff, *Helicobacter pylori accumulates photoactive porphyrins and is killed by visible light,* Antimicrob. Agents Chemother. **49**, 2822 (2005).

[31] A. Kumar, V. Ghate, M.-J. Kim, W. Zhou, G. H. Khoo, and H.-G. Yuk, *Kinetics of bacterial inactivation by 405 nm and 520 nm light emitting diodes and the role of endogenous coproporphyrin on bacterial susceptibility,* J. Photochem. Photobiol., B **149**, 37 (2015).

[32] D. E. Taupp, M. Nimtz, R. G. Berger, and H. Zorn, *Stress response of nidula niveotomentosa to uv-a light,* Mycologia **100** (2008).

[33] J. Vicente-Tejedor, M. Marchena, L. Ramírez, D. García-Ayuso, V. Gómez-Vicente, C. Sánchez-Ramos, P. de la Villa, and F. Germain, *Removal of the blue component of light significantly decreases retinal damage after high intensity exposure,* PloS one **13**, e0194218 (2018).

3

DATA DRIVEN MODELING OF THE REACTIVE OXYGEN SPECIES STIMULATED BY PHOTON ENERGY IN LIGHT THERAPIES

Light therapies can be used to treat fungal infections. A general mechanism is attributed to the generation of cytotoxic reactive oxygen species (ROS) due to light stimulation. The effectiveness of these therapies has been widely studied in the literature via conducting biological experiments, where fungi are exposed to light with various wavelengths and power. However, despite the large amount of work reporting the experimental results, few efforts have been given to build a mathematical model that describes the amount of generated ROS as a function of the photon energy and power of the stimulating light. The lack of such a model still hinders the optimization of the light doses. In this work, we propose a novel modeling method based on experimental data, so as to establish a mathematical relationship between the ROS concentration and the stimulating photon energy and light fluence (energy density). The anti-fungal experiments were performed on Candida albicans (C. albicans) using four LED light sources with different wavelengths ranging from 385nm to 450nm. Both the viability of the fungi and the ROS concentration therein were measured during the experiments. High fitting accuracy has been achieved by the model, which therefore demonstrates the effectiveness of the proposed modeling techniques.

3.1. INTRODUCTION

Light therapies can kill fungi and hence treat fungal infections. One of the major infectious types of fungi is *C. albicans* [1], which is widely found in nature, and commonly occurs as a superficial infection on mucous membranes, e.g. in mouths [2], vaginas [3]

and intestines [4]. The treatment of fungal infections by light is a non-antibiotic approach, and can avoid many side effects of antibiotic treatments, e.g. drug resistance [5].

Various blue light within the range of 400–470nm has been studied for anti-fungal therapies. The range with the most effective anti-fungal effect has been found in various studies to be 402–420nm [6]. For instance, 405nm light was proven to be highly effective against the pre-germinated spores of eight different types of fungi [7]. Besides, 415nm blue light was successfully applied to eliminate *C. albicans* in both in vitro and in vivo experiments [8].

The underlying mechanism of the anti-fungal effect of light has not yet been fully understood. A widely accepted hypothesis is that the photons from blue light can excite endogenous intracellular photosensitizers (PS), which in turn produces highly toxic ROS to cells [6], such as singlet oxygen ($^1O_2$), hydroxyl radials ($HO^\bullet$) and etc. Modeling the amount of generated ROS is hence an important issue in order to design the doses, e.g. the photon energy and light power, for an optimal effect in these therapies. However, developing such a mathematical model is a challenging task. The reason can be attributed to the complexities in the photosensitized oxidation reactions, in which the exact types of reactions that take place in the fungi and the exact types of ROS that are generated are actually uncertain [9].

ROS modeling has also been reported as an important problem, and studied in other therapies. For instance, modeling the ROS density in the plasma generated by dielectric barrier discharge has recently been reported in [10], which can be used in wound healing and dermatological therapies. Moreover, in photodynamic therapies (PDTs), i.e. the light therapies that apply exogenous PS, modeling the dynamic changes in ROS concentration has been well studied, e.g. [11–13]. For instance, a set of coupled differential equations are used to describe a PDT process, including seven Michaelis-Menten type equations [14] describing the dynamic changes in the concentrations of respectively the ground, singlet and triplet state of PS, the singlet and triplet state of oxygen, superoxide anions ($O_2^{-\bullet}$), and finally the ROS acceptors excluding the photosensitizer molecules [13]. Moreover, Monte Carlo simulations have been combined with these kinds of kinetic models, and proven to be an effective method for simulating light transport in biological tissues [15, 16]. Nevertheless, the aforementioned kinetic models demonstrate high nonlinearity, and moreover contain many different unknown parameters related to the PS characteristics that shall be determined from dedicated experiments. For instance, up to 21 parameters are required to describe the process of using some FDA or EMA approved PS [13].

Although anti-fungal light therapies are also believed to be caused by the PS that naturally exist in fungal cells, due to the aforementioned uncertain mechanisms of these processes (i.e. what types of reactions and produced ROS), modeling by the first principles becomes even more challenging than modeling a PDT process. To deal with these modeling challenges, a data-driven modeling approach is proposed in this work, which builds the functional relationship between the ROS concentration and two important parameters of the stimulating light, i.e. its photon energy and fluence, by fitting a parameterized model from experimental data. The model structure is motivated from the trend of the time sequence of $^1O_2$ concentrations at a time scale longer than one minute

Table 3.1: Main parameters used in the model

Quantity	*Unit*	*Symbol*
irradiance	mW/cm^2	E_e
fluence	J/cm^2	H_e
photon energy	eV	E_p
wavelength	nm	λ
ROS concentration	μM	y
reaction rate	none	k

as shown in [12], which is monotonically increasing in a fashion similar to the step response of a first order linear dynamic system [17]. Then, the reaction rate constant in this model is fitted to the photon energy of four different light wavelengths of 385nm, 405nm, 415nm and 450nm. It is also worth mentioning here that the complicated nonlinear dynamics of the first-principle PDT models mainly occur at a tiny time scale, i.e. below one second [12]. Fortunately, this transient behaviour is not of key importance for quantifying the long-term ROS accumulation in light therapies. In fact, many other studies have reported similar gradually increasing ROS accumulations in cells by light stimulation in a time duration of up to hundreds of minutes, e.g. [18, 19].

The contributions of this work are four folds. Firstly, the anti-fungal experiments on *C. albicans* were conducted using four LED light sources with four different wavelengths, including 385nm, 405nm, 415nm and 450nm. Both the viability of the fungi and the ROS concentrations were measured during the experiments. Secondly, a first order linear dynamic model is parameterized for the ROS variations, whose parameters were then estimated from the experimental data. Thirdly, the functional relationship between the reaction rate constants in the four fitted models respectively of 385nm, 405nm, 415nm and 450nm and their corresponding photon energy were constructed. Finally, a complete mathematical model of the ROS concentration induced by light irradiation is established, taking as variables the photon energy and fluence of the light. To the best of our knowledge, it is the first attempt to build such a mathematical model to mathematically describe the induced ROS in vitro without utilizing any exogenous PS. The main symbols used throughout the paper are defined in Table 3.1.

3.2. MATERIALS AND METHODS

3.2.1. LED LIGHT SOURCE DESIGN

Four different types of LEDs with the specified peak wavelengths respectively at 385nm, 405nm, 415nm and 450nm were applied in this work. The types of these LEDs and their main parameters are listed in Table 3.2. Their spectral power density (SPD) curves, as measured by a Maya2000Pro spectrometer (Ocean Optics, US), are depicted in Fig. 3.1(a). In this figure, every SPD curve is normalized with respect to its integral over the range of the measured wavelength, i.e. with each normalized SPD curve integrating to 1.

The LED light sources were designed following the procedures in [20], and were driven by a constant current source with PWM current level control to stabilize the output irradiance [21, 22]. The LED chips are arranged in either a 1.5cm-by-1.5cm square as a 4-by-

Table 3.2: Applied LEDs and their main parameters, where "W.L." stands for wavelength.

Specified Peak W.L. (nm)	*Type*	*Measured Peak W.L. (nm)*	*Photon Energy (eV)*
385nm	Vishay VLMU3500-385-120	386.2	3.21
405nm	Kingbright ATDS3534UV405B	400.9	3.09
415nm	LUMILEDS LUXEON LHUV-0415-A070	416.5	2.98
450nm	Cree XLampXPE2	447.6	2.77

4 array (385nm, 405nm and 415nm) or a 1cm-by-1cm square as a 3-by-3 array (450nm). The four LED light sources can deliver an irradiance of $50 mW/cm^2$ uniformly within a $60 cm$-diameter circle. Fig. 3.1(c) shows the simulated irradiance distribution, where the average irradiance in the 6cm-diameter circle is $49.39 mW/cm^2$, with a relative variation of only 6.56%. The irradiance was measured and confirmed by a PM100D power meter with a S120VC probe (Thorlabs, Newton, NJ, US). The schematic diagram of the electrical drive and control system and the experimental setup are illustrated in Figs. 3.1(b) and 3.1(d).

3.2.2. EXPERIMENTAL METHODS

ANTI-FUNGAL ASSAY

The *Candida albicans* strain used in this work is the 3147 (IFO 1594) strain (ATCC, US). The fungal viability was estimated by colony counting in terms of colony-forming units (CFU). After the concentration of the fungi reached 10^7 cells/ml, the fungal suspension was diluted by 10^4-fold with sterilized water. The diluted suspension was then spread on tryptic soy agar (TSA), and divided into a control group and a treatment group. The treatment group was respectively irradiated by one of the four LED light sources; while the control group was kept in the dark. The distances from the light sources to the fungi were around 10cm in all the cases to achieve the target irradiance of $50 mW/cm^2$.

More specifically, the light irradiation experiments were performed at each sampling instant of respectively 0, 5, 10, 15, 20, 25 and 30 minutes under the 385nm, 405nm and 415nm light exposure; and at 0, 15, 30, 45, 60, 75, 90, 105 and 120 minutes under the 450nm light irradiation, because it takes longer time for this wavelength to achieve a significant fungal inhibition at the same irradiance. At each of these sampling instants, one agar plate spread with the fungal suspension was exposed to the light; and another plate spread with the same suspension was kept in the dark. Moreover, only one plate was exposed to the light at each light irradiation experiment. Therefore, three plates were separately irradiated by the light at different time for each sampling instant. Afterwards, all the six plates were cultured for 24-48 hours at 26^oC, before the colony counting. These steps resulted in triplicate experiments for each treatment time interval, whose raw CFU data were then processed to produce the mean and standard deviation of the viability

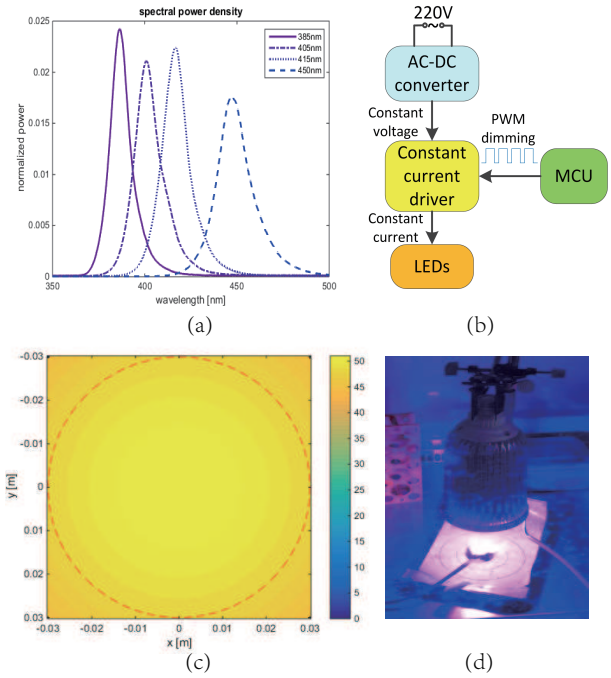


Figure 3.1: LED light source design: (a) normalized SPD of the four types of LEDs, (b) electrical scheme, (c) simulated irradiance distribution, (d) photo of the experimental setup.

rates at each sampling instant. The significance of the growth inhibition of the fungi before and after the light treatment was tested by the Student's t-test.

ROS ASSAY

For the homogeneity of the ROS measurements, the concentrations of the fungi used in all the ROS assays were controlled within the same range by measuring its absorption of 450nm light with a U-3900H spectrophotometer (Hitachi, Japan). More specifically, the absorption levels measured by this equipment were always controlled in the range of 6.5-7. Then, the fungi were centrifuged and separated from the medium, and were dissolved in a 500-fold dilution of the ROS fluorescent probe (DCFH-DA assay kit, Beyotime Institute of Biotechnology, China) by phosphate buffer saline (PBS). After incubated at 37^oC in a shaker for half an hour, the suspension was centrifuged for three times to remove the redundant probe. Then, the suspension was seeded into a 96-well plate, and divided into a treatment group and a control group. The treatment group was irradiated respectively by the four LED light sources; while the control group was kept in the dark. Immediately after the light treatment, the intracellular ROS levels were measured as the fluorescent levels in a VL0L0TD0 Varioskan LUX microplate reader (Thermo Fisher, US), with the exciting and emitting wavelength respectively set at 488nm and 525nm. For each treatment interval, the plates were continuously exposed to the light, and were then discarded after the measurement by the microplate reader.

Triplicate experiments were performed in a similar fashion as described in the anti-fungal assays; i.e. respectively at the sampling instant of 0, 5, 10, 15, 20, 25 and 30 minutes under the 385nm, 405nm and 415nm light exposure; and at 0, 15, 30, 45, 60, 75, 90, 105 and 120 minutes under the 450nm light irradiation. The raw data were processed to produce the mean and standard deviation of the ROS fluorescent levels for each treatment time interval.

3.2.3. DYNAMIC MODEL OF ROS CONCENTRATIONS INDUCED BY PHOTON ENERGY

As introduced in Sec. 3.1, although comprehensive models of the photosensitized oxidation reactions in PDTs have already been established in the literature, models of the induced ROS in vitro without utilizing any exogenous PS are not yet available. The difficulties can be attributed firstly to the many unknown parameters that must be determined from dedicated experiments, and secondly to the lack of knowledge about what types of reactions, endogenous PS and oxidants are exactly involved in anti-fungal therapies. To avoid these difficulties and reduce the experimental burdens, we propose to use a reduced model structure and estimate its parameters from the ROS concentration data observed from experiments. The reduction from the highly coupled nonlinear Michaelis-Menten type models [13] does not take the short-time transient behaviour into account, and is only valid for modeling the ROS accumulations at a time scale larger than one minute. This is motivated by the fact that at a macroscopic time scale, the increasing trend of the singlet oxygen concentration [12] is similar to the step response of a first order linear dynamic system [17].

Denote the ROS concentration by y. The first order linear system description of its dynamics can be expressed by the following differential equation.

$$\frac{dy(t)}{dt} = -k \cdot y(t) + u, \tag{3.1}$$

where $t \geq 0$ is the time instant; $k > 0$ is the reaction rate constant; and $u > 0$ represents an unknown input, and is constant in the case of a step response. This model is actually an integrator, and is thus suitable to describe the accumulation of ROS over time. The solution of Eq. (3.1) takes an analytic form [17] as

$$\begin{aligned} y(t) &= e^{-kt} \cdot y_0 + \int_0^t u \cdot e^{-k(t-\tau)} d\tau, \\ &= \frac{u}{k} + \left(y_0 - \frac{u}{k}\right) \cdot e^{-kt}, \end{aligned}$$

where y_0 denotes the initial value of y at time instant 0. When $u > 0$ is constant for $t \geq 0$, this equation defines the step response of the dynamic system defined by Eq. (3.1). For brevity, the time variable will be omitted in what follows.

Equivalently, one can parameterize u as $u = k \cdot (y_0 + u')$, with $u' \geq 0$ arbitrary. Then, the above equation is further simplified to

$$y = y_0 + u' \cdot \left(1 - e^{-kt}\right).$$

Bringing y_0 to the left of the equal sign and dividing both sides by y_0, one can obtain the following equation of the relative change of the ROS concentration with respect to its initial value.

$$\underbrace{\frac{y-y_0}{y_0}}_{x} = \underbrace{\frac{u'}{y_0}}_{r} \cdot \left(1-e^{-kt}\right). \tag{3.2}$$

For brevity, denote $\frac{y-y_0}{y_0}$ by x and $\frac{u'}{y_0}$ by r in what follows.

Note that since the ROS concentrations are actually measured with fluorescent probes, by e.g. a microplate reader, it is more meaningful to take the relative value x in Eq. (3.2). Besides, it is well known that ROS naturally exists in cells and regulates a great number of biochemical reactions, which generally accounts for 2% of the total oxygen consumed by mitochondria under a "normal" condition [23]. Therefore, an initial ROS concentration in the fungi always exists, and contributes to a nontrivial initial fluorescent level, i.e. $y_0 > 0$, before being irradiated by light.

Obviously, Eq. (3.2) is a monotonically increasing function of t. The only parameters of this model to be estimated are k and r. The formulation of Eq. (3.2) indicates that r can be interpreted as the ratio of the initial total concentration of the other molecules involved in the reactions (e.g. the triplet oxygen and all the three states of the endogenous PS) to the initial concentration of the ROS.

3.2.4. FUNCTIONAL RELATIONSHIP BETWEEN REACTION RATES AND PHOTON ENERGY

In classical PDT models, the PDT dose is defined as the number of photons absorbed by the PS, and is related to the irradiance and the photon energy [12, 24]. The kinetic PDT equations are henceforth parameterized by the irradiance and photon energy of the stimulating light.

The energy of a photon is inversely proportional to the wavelength of the light, which is usually given in the unit of electron-volt (eV, and $1eV = 1.602 \times 10^{-19} J$), i.e.

$$E_p = \frac{h \cdot c}{\lambda} = \frac{1.24}{\lambda}, \tag{3.3}$$

where h and c are respectively the Planck's constant and the speed of light. Here, the unit of λ shall be converted from nanometers to microns. Photon energy will also be used to quantify a specific light source in what follows.

For in vitro anti-fungal experiments, the photon energy is directly absorbed by fungal cells instead of tissues. On the other hand, according to the experimental results to be presented later in Sec. 3.3, the reaction rate constants in Eq. (3.2) estimated from the experimental data using the four light spectra are all different, when keeping the light irradiance at the same level. More specifically, the rate constants at 415nm and 450nm are respectively the largest and the smallest; while those at 385nm and 405nm first show a "roll-off" after the rate peak near 415nm, and are then followed by an increasing trend again (see Table 3.3 and Fig. 3.6). This pattern actually corresponds well to the more reported practices of using 400-420nm light in the anti-fungal experiments. Motivated by this observation, one can describe this dependence of the reaction rate constant on

$$k = \begin{cases} a + b \cdot \sin(d \cdot E_p + \varphi), & E_p \le \bar{E} \\ a + b \cdot \exp\left\{-[\alpha \cdot (E_p - \mu)]^c\right\} \cdot \sin(d \cdot E_p + \varphi), & E_p > \bar{E} \end{cases} \tag{3.4}$$

$$x(E_p, H_e) = \begin{cases} r \cdot \left\{1 - \exp\left\{-\frac{1}{E_e} \cdot \left[a + b \cdot \sin(d \cdot E_p + \varphi)\right] \cdot H_e\right\}\right\}, & E_p \le \bar{E} \\ r \cdot \left\{1 - \exp\left\{-\frac{1}{E_e} \cdot \left\{a + b \cdot e^{-[\alpha \cdot (E_p - \mu)]^c} \cdot \sin(d \cdot E_p + \varphi)\right\} \cdot H_e\right\}\right\}, & E_p > \bar{E} \end{cases} \tag{3.5}$$

the photon energy by Eq. (3.4). Here, $a, b, c, d, \alpha, \mu, \varphi$ are the parameters to be estimated; and $\bar{E}$ is the point at which the reaction rate starts to roll off. In the model (3.4), the sinusoidal function describes the peak near 415nm and the decreasing trend for longer wavelengths up to 460nm; while the exponentially decaying term, $\exp\left\{-[\alpha \cdot (E_p - \mu)]^c\right\}$, is to account for the roll-off after the rate peak.

3.2.5. MODELING ROS CONCENTRATION AS A FUNCTION OF PHOTON ENERGY AND FLUENCE

Now, by substituting Eq. (3.4) into Eq. (3.2) and noting that $t = H_e/E_e$, the model of the relative ROS concentration x as a function of the photon energy and fluence of the stimulating light can be finally derived, which takes $a, b, c, d, r, \alpha, \mu, \varphi$ and the light irradiance E_e as parameters. The model takes the form of Eq. (3.5).

3.2.6. MODELING FUNGAL VIABILITY AS A FUNCTION OF PHOTON ENERGY AND FLUENCE

Although a viability model is not required in the aforementioned ROS model, it is also relevant to further understand the effect of the photon energy and fluence on the efficiency of eliminating the fungi by the light. Such a mathematical relationship can be built in a similar fashion as in building the ROS model.

To this end, one first needs to fit a time-varying viability model of the fungi when being exposed to the light of each wavelength. Since it takes some time for the ROS concentration to reach a sufficient level to kill the fungi, a piecewise function including a "shoulder" [25] is suitable to describe such a process, i.e.

$$\rho(t) = \begin{cases} 1, & t \le \tau \\ e^{-k_v(t-\tau)}, & t > \tau \end{cases} \tag{3.6}$$

where $\rho(t)$ is the survival rate at time t; k_v is the decaying rate coefficient; and τ is the time constant when the inactivation starts.

The next step is also to fit the parameters k_v and τ to the photon energy, because they vary among the four different light wavelengths. Also according to the experimental results to be presented later in Sec. 3.3, the models of k_v and τ are respectively parameterized as follows.

$$k_v = \exp\left(p_1 \cdot E_p^4 + p_2 \cdot E_p^3 + p_3 \cdot E_p^2 + p_4\right), \tag{3.7}$$

$$\tau = q_1 \cdot E_p^3 + q_2 \cdot E_p^2 + q_3 \cdot E_p + q_4. \tag{3.8}$$

Here, $p_i, q_i, i = 1,2,3,4$ are the coefficients to be estimated. The exponential function in Eq. (3.7) is to enforce $k_v > 0$, since a pure polynomial function cannot ensure this, when k_v is close to zero.

Similarly, by substituting Eqs. (3.7, 3.8) into Eq. (3.6) and taking $t = H_e/E_e$, the model can be finally written as Eq. (3.9).

$$\rho(E_p, H_e) = \begin{cases} 1 & , \quad H_e \le E_e \cdot \sum_{i=1}^{4} q_i \cdot E_p^{4-i} \\ \exp\left\{-\exp\left(\sum_{i=1}^{3} p_i \cdot E_p^{5-i} + p_4\right) \cdot \left(\frac{H_e}{E_e} - \sum_{i=1}^{4} q_i \cdot E_p^{4-i}\right)\right\} & , \quad \text{otherwise} \end{cases} . \tag{3.9}$$

3.3. RESULTS

3.3.1. VIABILITY MEASUREMENTS OF C. ALBICANS

In the experiments with the 385nm, 405nm and 415nm light sources, no significant growth inhibition of *C. albicans* was observed until being irradiated for ten minutes ($p < 0.05$). After being exposed to the light for 25 or 30 minutes, the survival rates of the fungi all dropped below 20%. On the other hand, in the experiment with the 450nm light source, significant elimination of *C. albicans* was observed after being irradiated for 100 minutes ($p < 0.05$); while after two hours, about 80% of the fungi were inhibited. The viability rates of the four experiments are depicted in Fig. 3.2. As an illustrative example, the fungal growth on the agar plates after being irradiated by the 385nm light for different durations from 0 to 30 minutes is shown in Fig. 3.2(e). The decreasing trend of the CFU can be clearly observed.

3.3.2. ROS MEASUREMENTS

The time sequences of the measured ROS concentrations in the *C. albicans* were measured from the experiments using all the four different light sources. The measured fluorescent levels y were processed according to the definition of Eq. (3.2) as $x = \frac{y-y_0}{y_0}$, where y_0 is the initial level. The relative changes in the ROS concentrations due to the light stimulation are plotted in Fig. 3.3.

To visualize the gradual accumulations of the ROS in the fungi as the light exposure continues, the time-lapse images of the fungal cells after being irradiated by the 405nm light for different durations from 0 to 30 minutes were taken by an Axio Observer A1 inverted fluorescence microscope (ZEISS, Germany), and are illustrated in Fig. 3.3(e). The exciting and emitting wavelength were set respectively at 488nm and 525nm. The increasing trend of the ROS concentrations can be clearly observed.

3.3.3. ESTIMATION OF THE PARAMETERS OF EQ. (2)

Eq. (3.2) contains two parameters to be estimated, i.e. r and k. By definition, this estimation problem is nonlinear, and can be solved by a standard nonlinear least-squares (NLS) algorithm [26], e.g. the Levenberg-Marquardt method. The dependence of the RMSE fitting errors on the values of r and k is shown in Fig. 3.4. It can be seen that in all the four cases, the optimum of the parameter pair (r, k) lies on the flat bottom of a

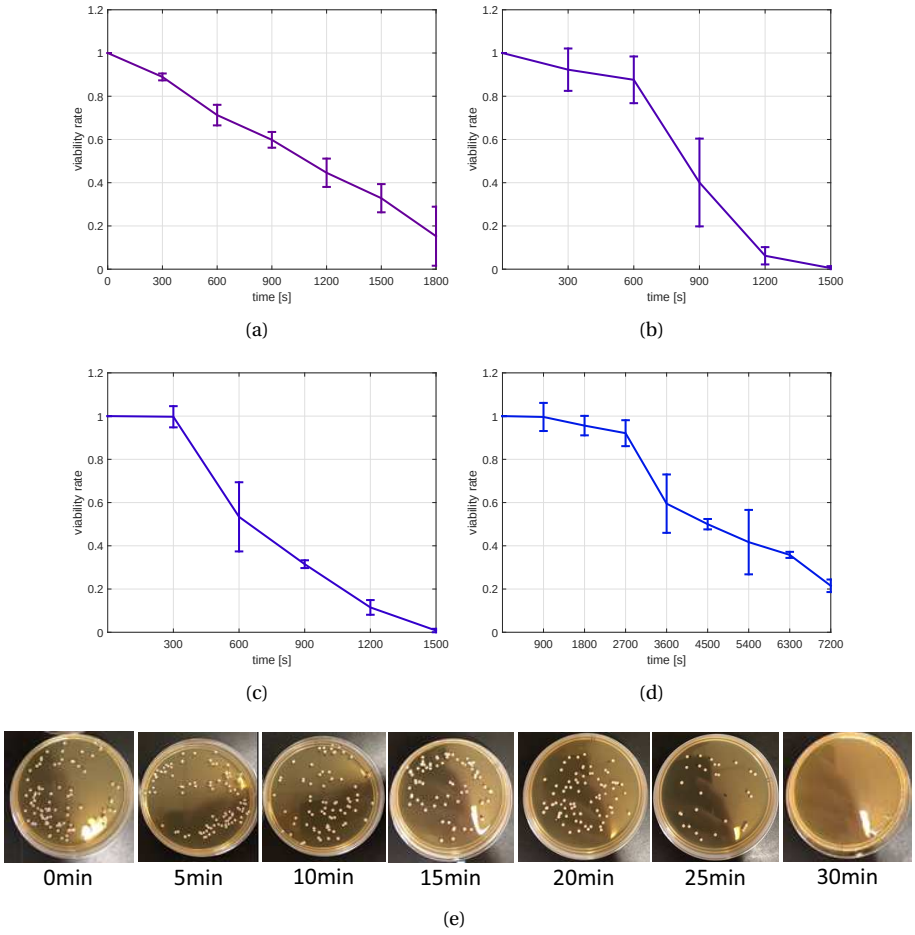


Figure 3.2: Measured viability rates of the *C. albicans* irradiated respectively by (a) 385nm, (b) 405nm, (c) 415nm, (d) 450nm LED light source, and (e) the pictures of the fungal growth on the agar plates after being irradiated by the 385nm light for different durations from 0 to 30 minutes.

narrow valley roughly within the range of $30 \leq r \leq 3000$ and $10^{-4} \leq k \leq 10^{-6}$. The gradients of the fitting errors with respect to this pair are approximately zero within this valley. In other words, the NLS optimization is likely to terminate at any point in this valley depending on the specified initial values and stopping conditions, which can also been seen from Fig. 3.4. Based on this observation, it is better to fix r to a value in the range of $30 \leq r \leq 3000$, and estimate k by linear least-squares (LS), which is guaranteed to result in a unique global optimal solution. Moreover, it is also necessary to set r to a same value, in order to compare the effects of the four wavelengths on the reaction rates k.

On the other hand, note that the fungi used in this work are of the same type, and

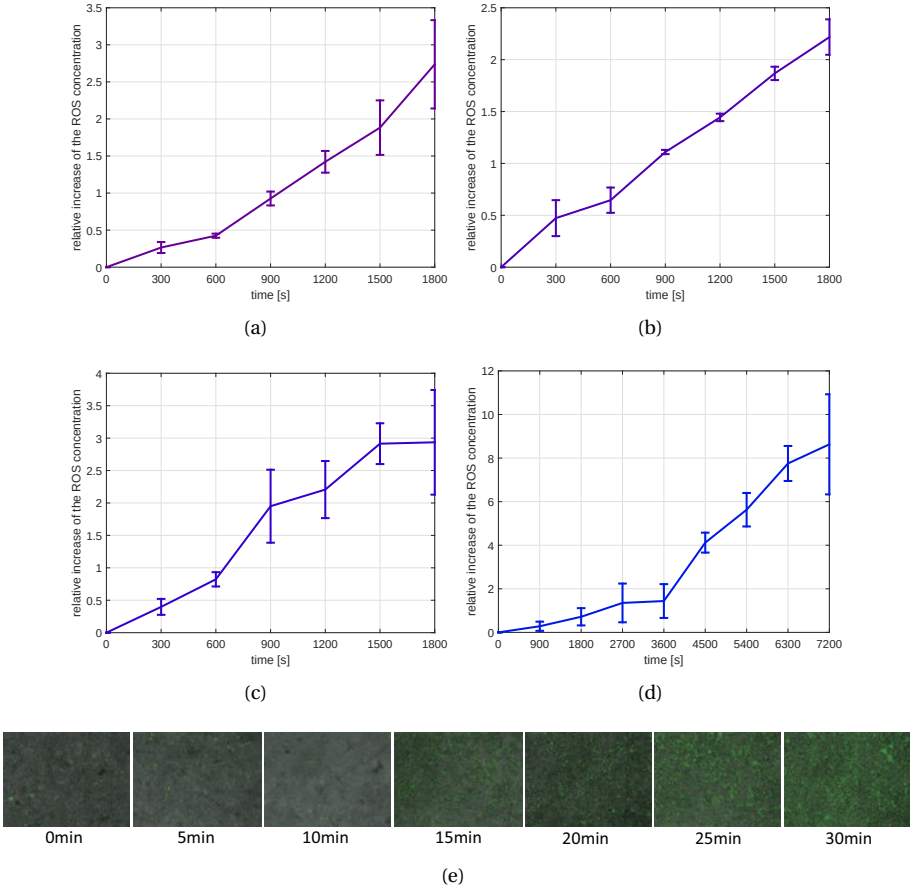


Figure 3.3: Measured relative changes in the ROS fluorescent levels in the *C. albicans* irradiated respectively by (a) 385nm, (b) 405nm, (c) 415nm, (d) 450nm LED light source, and (e) the time-lapse ROS fluorescent images of the fungal cells taken from 0 to 30 minutes with a 5min time difference.

were cultivated and processed following exactly the same protocols. It is hence also reasonable to assume that the parameter r in all the experiments is of the same value. According to the interpretation of r in Eq. (3.2), this physically means that the ratios of the initial total concentration of the other molecules involved in the reactions (e.g. the triplet oxygen and all the three states of the endogenous PS) to the initial ROS concentration can be assumed to be at the same level in all the experiments. When there is an ample amount of oxygen molecules in the cells, i.e. when r is dominated by the ratio between the concentration of the initial triplet oxygen and the initial ROS, one can take $r = 50$, since in the normal condition of mitochondrial respiration, ROS takes about 2% of the total oxygen consumption [23]. More specifically, by fixing r, k can be estimated

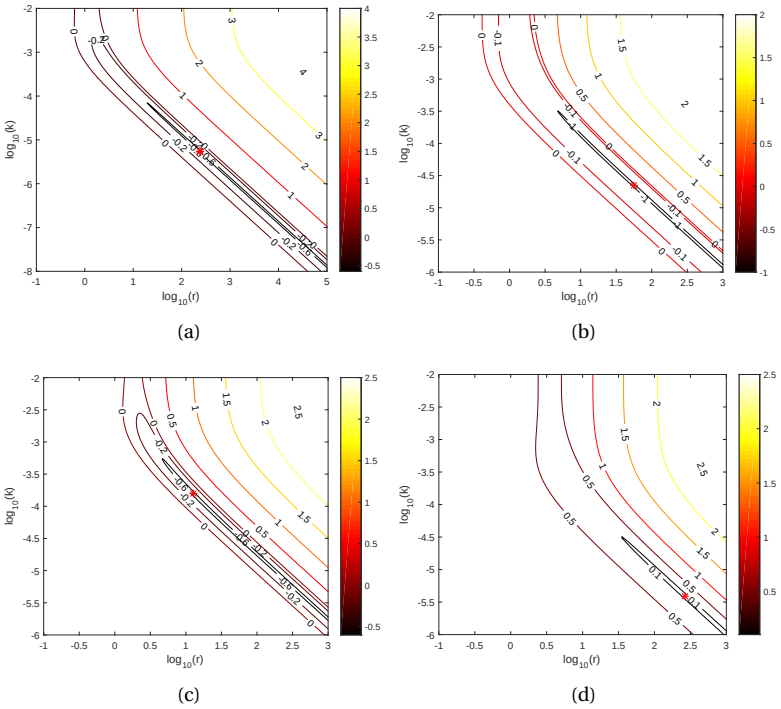


Figure 3.4: Dependence of the RMSE fitting errors on the values of r and k: (a) 385nm, (b) 405nm, (c) 415nm, (d) 450nm. Contours: $\log_{10}(RMSE)$; Stars: the NLS estimates.

by solving the following LS problem.

$$k \cdot t = -\ln\left(1 - \frac{x}{r}\right). \tag{3.10}$$

The LS estimates and the RMSE fitting errors are listed in Table 3.3. As an illustrative example, the model fitted to the data from the 415nm light irradiation experiment is depicted in Fig. 3.5.

3.3.4. ESTIMATION OF THE PARAMETERS OF EQ. (4)

The estimated reaction rate constants in Table 3.3 and their corresponding photon energy in Table 3.2 are plotted in Fig. 3.6. Note that in Eq. (3.4), there are totally seven unknown parameters to be estimated, i.e. $a, b, c, d, \alpha, \mu, \varphi$. However, there are only four data points, which can at most uniquely determine four parameters. To solve them, d was first determined by estimating the period of the sinusoidal function from the four target points, by noting that the range of the photon energy shall be within one complete period of the sinusoidal function (otherwise, there will be multiple peaks). On the other hand, $\bar{E}$ that determines the roll-off point shall be between 2.98 and 3.09; and α, μ are intended to shift and normalize the photo energy into the range between 0 and 1.

Table 3.3: Estimated reaction rate constants and RMSE fitting errors

	k	*RMSE*
385nm	2.65E-5	0.240
405nm	2.50E-5	0.056
415nm	3.68E-5	0.216
450nm	2.22E-5	1.233

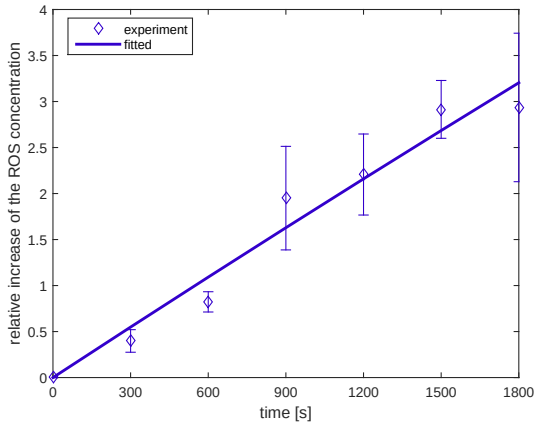


Figure 3.5: Fitted curve to the measured relative ROS concentration data from the 415nm light irradiation experiment.

This then reduces the number of unknown parameters to three, which can be solved by a standard NLS algorithm. The finally estimated parameters are listed in Table 3.4, which perfectly fits to the four target points with an RMSE fitting error of 7.86E-7.

3.3.5. SIMULATING THE ROS MODEL WITH VARIOUS PHOTON ENERGY AND FLUENCE

With the parameters in Table 3.4 and taking $r = 50$ and $E_e = 50mW/cm^2$, the model of the ROS concentration as a function of the stimulating photon energy and fluence, i.e. Eq. (3.5), was finally simulated. The results are plotted in Fig. 3.7.

Table 3.4: Estimated parameters of Eq. (3.4)

a	b	c	d	α	μ	φ	$\bar{E}$
2.66E-5	1.90E-5	3.025	13.57	3.100	2.667	-31.55	3.012

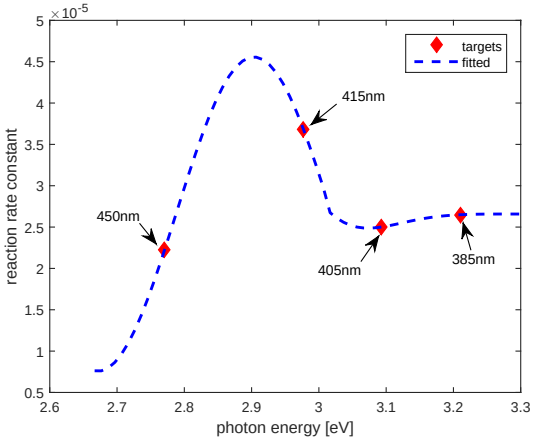


Figure 3.6: Fitted model of the reaction rate constants as a function of the photon energy. The arrows indicate the experimenting light sources. The simulated range of the photon energy $2.7 \sim 3.3eV$ corresponds to the wavelength range of $375 \sim 460nm$.

3.3.6. SUMMARY OF THE PROCEDURES TO BUILD THE ROS MODEL

Although a single strain of *C. albicans* was tested in this work, the proposed modeling method can actually be applied to other fungal species, as well. The experimental and modeling procedures follow the steps shown in Fig. 3.8.

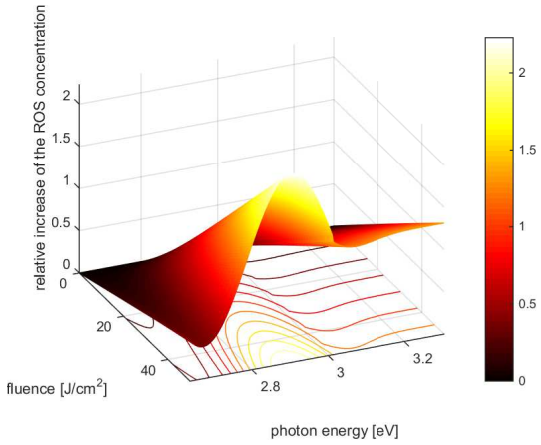
3.3.7. FITTING AND SIMULATION OF THE VIABILITY MODEL

Since the main purpose of this work is to build the ROS model, and also for brevity, the detailed parameter estimation procedures for the viability model will be omitted; and only the fitting results are presented here.

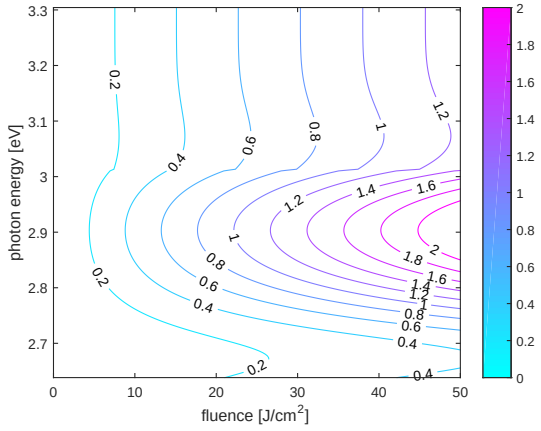
The parameters k_v and τ were first fitted to the viability data as shown in Fig. 3.2, whose estimated values for all the four wavelengths are listed in Table 3.5. The estimated coefficients of the polynomials in Eqs. (3.7, 3.8) are listed in Table 3.6, which resulted in the fitted curves as illustrated in Fig. 3.9. With these estimated parameters, the viability model was finally simulated. The results are plotted in Fig. 3.10.

Table 3.5: Estimated parameters of Eq. (3.6)

	k_v	τ *(in sec.)*	*RMSE*
385nm	1E-3	226.80	0.235
405nm	7.06E-3	770.70	0.056
415nm	3.85E-3	447.30	0.211
450nm	2.89E-4	2232.72	1.176



(a)



(b)

Figure 3.7: Relative increase of the ROS concentration in response to various photon energy and fluence: (a) 3D plot, (b) contour plot.

Table 3.6: Estimated coefficients of the polynomials in Eqs. (3.7, 3.8)

	p_1 or q_1	p_2 or q_2	p_3 or q_3	p_4 or q_4	*RMSE*
Eq. (3.7)	-37.23	287.19	-620.59	841.56	9.20E-16
Eq. (3.8)	-1.53E5	1.39E6	-4.19E6	4.22E6	2.38E-9

3.4. DISCUSSION

In this study, anti-fungal light irradiation experiments were conducted using four different wavelengths, in which both the viability of the fungi and the generated ROS therein

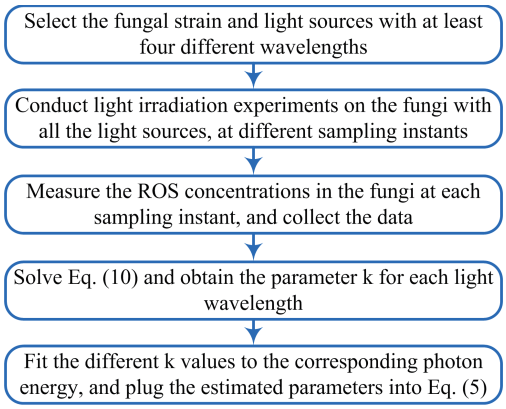
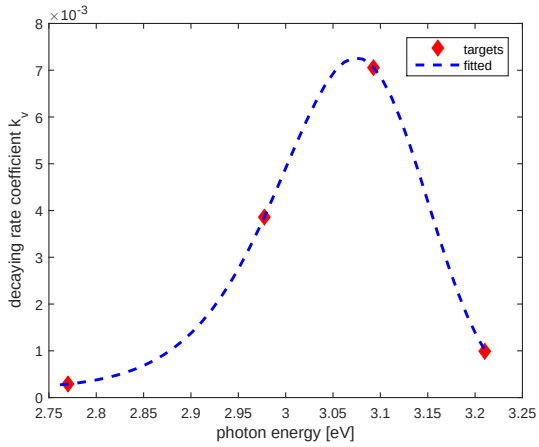


Figure 3.8: Procedures of implementing the data-driven modeling method.

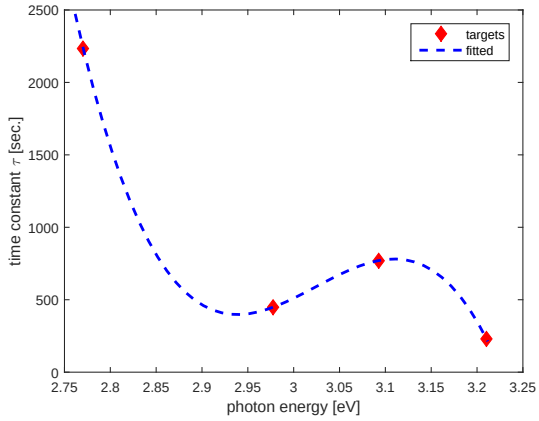
were measured. The main objectives are to compare the effects of four different wavelengths in terms of their effect in inducing ROS, and to build a mathematical model that relates the ROS generation with the photon energy and the fluence of the stimulating light.

The viability data in Fig. 3.2 show that the 415nm light performed best in eliminating *C. albicans* in terms of the lowest survival rate of the fungi after having been irradiated for 25 minutes. The longest wavelength 450nm turned out to be the worst in this aspect, because of the much longer time it took to significantly kill the fungi. On the other hand, a "shoulder" in the viability curve of the fungi can be clearly seen in most cases. The shoulders indicate that the fungi were not immediately killed right after being exposed to the light. This can be attributed to the fact that the main hypothesized mechanism of anti-fungal light therapies is the cytotoxicity due to the induced ROS from the interaction between the light and the endogenous PS. To generate enough toxicity to kill the fungi, the ROS in the cells needs to accumulate to a sufficient level. The turning point appeared when all the four wavelengths led to more than 40% fungal elimination. These points were respectively at 900sec for 385nm and 405nm, at 600sec for 415nm, and at 3600sec for 450nm. At these points, the relative ROS increases were of similar values around 1, i.e. $1.11 \pm 0.02, 0.93 \pm 0.09, 0.82 \pm 0.11, 1.44 \pm 0.78$ respectively for the 385nm, 405nm, 415nm and 450nm light. On the other hand, in the case of the 385nm-light irradiation, the fungi appeared to start degenerating sooner than the cases of using the 405nm and 415nm light. This indicates that the UVA light of 385nm may cause other more dramatic inhibiting effects to the fungi besides inducing ROS, e.g. causing tryptophan photodegradation within the cells [27]. Moreover, as can be observed from the simulated fungal viability rates in response to various photon energy and fluence as in Fig. 3.10, the most effective anti-fungal photon energy is in the range of 2.85~3.2eV, corresponding to wavelength range of 387.5~435nm.

Due to the aforementioned challenges to parameterize the model of the induced ROS concentrations in vitro without utilizing any exogenous PS, a data-driven modeling approach was proposed. This modeling approach mainly takes into account the long time



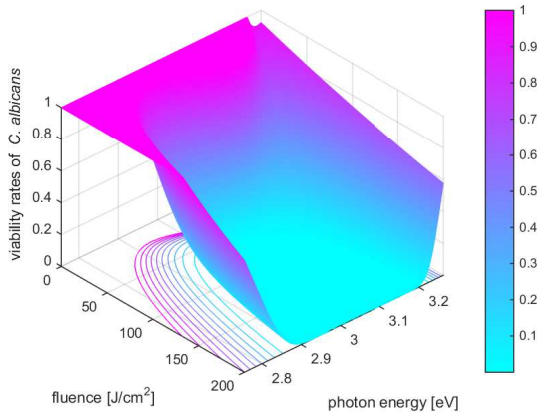
(a)



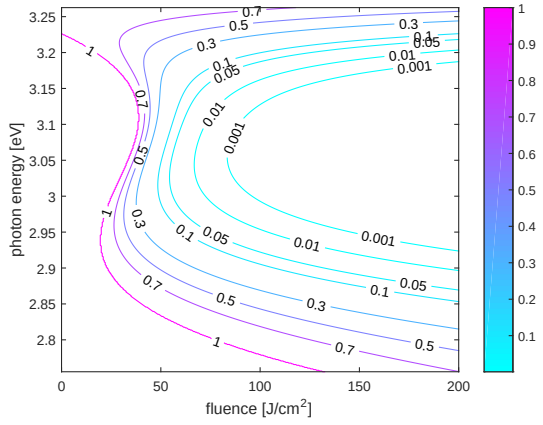
(b)

Figure 3.9: Fitted model of k_v and τ as a function of the photon energy: (a) k_v as in Eq. (3.7), (b) τ as in Eq. (3.8).

scale ROS accumulating effect during light irradiation. The analysis of the dependence of the RMSE fitting errors on the values of the two model parameters, i.e. r and k, show that their optimal values are actually trapped within a flat and narrow valley. In this valley, the gradients of the cost function with respect to these two parameters are almost trivial. Therefore, instead of leaving the optimization to stop at a random point in this valley (depending on the randomly given initial value and stopping criteria), the value of r was fixed to 50. Consequently, the reaction rate constant k can be estimated as a linear LS problem, which ensures that the unique global optimal value of k can be found. Here, a mild assumption is made that r is dominated by the ratio between the concentrations of the initial triplet oxygen and the initial ROS. Therefore, its value can be taken as 50,



(a)



(b)

Figure 3.10: Variation of the fungal viability rates in response to various photon energy and fluence: (a) 3D plot, (b) contour plot.

due to the aforementioned fact that in the normal condition of mitochondrial respiration, ROS takes about 2% of the total oxygen consumption. Good agreement of the fitted functions with the experimental data was achieved for the four data sets, as shown in Table 3.3 and Fig. 3.5. This demonstrates the validity of the proposed model structure and the assumption.

The estimated reaction rate constants show a dependence on the wavelengths or their corresponding photon energy, which peak near 415nm and decline to the minimum as the wavelength increases to 460nm. According to Eq. (3.2), the reaction rate constant determines the speed of generating ROS in the photochemical reaction. The larger the value, the faster the accumulation of the cytotoxic ROS, and thus the faster the elimination of the fungi. On the other hand, the roll-off effect of the reaction rate

at even shorter wavelengths, i.e. 385nm and 405nm, indicates that the sensitivity of the endogenous PS to UVA light is lower than that to the light in the range of 415~426nm. However, the similar anti-fungal effect of 385nm and 405nm with that of 415nm indicate that other inhibiting effects to the fungi besides inducing ROS might have also been induced by these wavelengths, e.g. causing tryptophan photodegradation [27].

It is known in the literature that there are several types of endogenous PS responsible for anti-fungal effects, including various kinds of porphyrins and flavins, depending on the fungal species. The Soret bands of porphyrins are generally in the range of 400~410nm [28]; while some types can reach 413~416nm [29]. On the other hand, the Soret bands of flavins are usually at longer wavelengths. For instance, the peak absorption of acriflavin neutral (or euflavine) is at 436nm [30]; and that of cytochrome-flavin complex is at 427nm [31]. In fact, the fitted model of Eq. (3.4) reveals that in the experiments the peak absorption wavelength is at 426.6nm or 2.907eV in photon energy. This indicates that there may be more than one type of endogenous PS in the $C.\ albicans$ studied in this work, which all contribute to the generation of ROS.

By simulating the model of the ROS concentration as a function of the photon energy and fluence, the following observations can be further made. First, when the fluence is small, i.e. $H_e < 5J/cm^2$, the induced ROS concentration is not much for any photon energy in the range of 2.7~3.3eV. This indicates that the radiometric energy of the stimulating light needs to be higher than a threshold to trigger sufficient ROS generation, no matter how large the photon energy is. Second, as the fluence accumulates to a certain level, i.e. $H_e > 10J/cm^2$, the effects of the photon energy start to manifest. For this studied fungal strain, the range of photon energy 2.85~2.95eV (or 420~435nm in wavelength) demonstrated a higher efficiency in generating ROS. These observations can be helpful to design the therapeutic devices and the doses used in treating the infections caused by this type of fungi.

On the other hand, the proposed model in this work also has some limitations. First, the proposed model structure is simplified, and cannot describe the transient dynamics of the PS and ROS. As another limitation, the proposed model does not distinguish different types of ROS. As being fitted to the general oxidative stress measured by the standard ROS assay kit, the output of the model is the changing rate of the total ROS accumulation, and hence contains the contributions from all the existing PS in the fungi.

3.5. CONCLUSIONS

In this work, a modeling approach has been developed to mathematically describe the induced ROS in fungi, as a function of the photon energy and fluence of the stimulating light. The method of estimating the model parameters from experimental data has also been proposed and verified. The fitting results agree well with the main trends of the experimental data at long time scales, e.g. from tens of minutes to a few hours. This indicates that the proposed model structures and the parameter estimation methods are effective to calculate the amount of accumulating ROS in the $C.\ albicans$, when being stimulated by the photon energy in the range of $2.7 \sim 3.3eV$.

As another main conclusion, the photon energy within the range of 2.85~2.95eV (or 420~435nm in wavelength) is more effective in generating ROS in the fungi studied in this work, and is hence more effective in treating the infections caused by this type of

fungi.

As a potential future extension, the data-driven modeling approach can be further extended to account for the individual contributions of the main types of endogenous PS in the fungi to the generation of ROS. Another extension will be to develop a data-driven approach to model the in vivo ROS generation.

REFERENCES

[1] R. Calderone and W. Fonzi, *Virulence factors of candida albicans,* Trends Microbiol. **9**, 327 (2001).

[2] T. Arendorf and D. Walker, *The prevalence and intra-oral distribution of candida albicans in man,* Arch. Oral Biol. **25**, 1 (1980).

[3] J. Sobel, *Vaginitis,* N. Engl. J. Med. **337**, 1896 (1997).

[4] A. Erdogan and S. Rao, *Small intestinal fungal overgrowth,* Curr. Gastroenterol. Rep. **17**, 16 (2015).

[5] T. Maisch, *A new strategy to destroy antibiotic resistant microorganisms: antimicrobial photodynamic treatment,* Mini-Rev. Med. Chem. **9**, 974 (2009).

[6] R. Yin, T. Dai, P. Avci, A. E. Jorge, W. de Melo, D. Vecchio, Y. Huang, A. Gupta, and M. Hamblin, *Light based anti-infectives: ultraviolet C irradiation, photodynamic therapy, blue light, and beyond,* Curr. Opin. Pharmacol. **13**, 731 (2013).

[7] W. Trzaska, H. Wrigley, J. Thwaite, and R. May, *Species-specific antifungal activity of blue light,* Sci. Rep. **7**, 4605 (2017).

[8] Y. Zhang, Y. Zhu, J. Chen, Y. Wang, M. Sherwood, C. Murray, M. Vrahas, D. Hooper, M. Hamblin, and T. Dai, *Antimicrobial blue light inactivation of candida albicans: In vitro and in vivo studies,* Virulence **7**, 536 (2016).

[9] M. Baptista, J. Cadet, P. D. Mascio, A. Ghogare, A. Greer, M. Hamblin, C. Lorente, S. Nunez, M. Ribeiro, A. Thomas, M. Vignoni, and T. Yoshimura, *Type I and II photosensitized oxidation reactions: Guidelines and mechanistic pathways,* Photochem. Photobiol. **93**, 912 (2017).

[10] X. Wang, Y. Liu, Z. Tan, and L. Chang, *Effects of oxygen concentration on the reactive oxygen species density under different operating conditions in atmospheric-pressure helium/oxygen pulsed dielectric barrier discharge,* IEEE Access **7**, 69748 (2019).

[11] K. Wang, S. Mitra, and T. Foster, *A comprehensive mathematical model of microscopic dose deposition in photodynamic therapy,* Med. Phys. **34**, 282 (2007).

[12] T. Zhu, J. Finlay, X. Zhou, and J. Li, *Macroscopic modeling of the singlet oxygen production during pdt,* in *Proc. SPIE*, Vol. 6427 (2007) p. 642708.

[13] M. Kim, A. Ghogare, A. Greer, and T. Zhu, *On the in vivo photochemical rate parameters for PDT reactive oxygen species modeling,* Phys. Med. Biol. **139**, R1 (2017).

[14] K. Johnson and R. Goody, *The original michaelis constant: Translation of the 1913 Michaelis-Menten paper,* Biochemistry **50**, 8264 (2011).

[15] G. Kareliotis, S. Liossi, and M. Makropoulou, *Assessment of singlet oxygen dosimetry concepts in photodynamic therapy through computational modeling,* Photodiagn. Photodyn. Ther. **21**, 224 (2018).

[16] K. Beeson, E. Parilov, M. Potasek, M. Kim, and T. Zhu, *Validation of combined Monte Carlo and photokinetic simulations for the outcome correlation analysis of benzoporphyrin derivative-mediated photodynamic therapy on mice,* J. Biomed. Opt. **24**, 1 (2019).

[17] K. Åström and B. Wittenmark, *Computer controlled systems: theory and design* (Prentice Hall, 1984).

[18] J. Zhang, D. Xing, and X. Gao, *Low-power laser irradiation activates Src tyrosine kinase through reactive oxygen species-mediated signaling pathway,* J. Cell. Physiol. **217**, 518 (2008).

[19] A. Lynnyk, M. Lunova, M. Jirsa, D. Egorova, A. Kulikov, Šárka Kubinová, O. Lunov, and A. Dejneka, *Manipulating the mitochondria activity in human hepatic cell line Huh7 by low-power laser irradiation,* Biomed. Opt. Express **9**, 1283 (2018).

[20] J. Dong and Z. Zhang, *Design of led light for stimulating cells in the study of light therapies,* in *Proc. 15th SSLChina: IFWS* (2018) pp. 74–77.

[21] J. Dong and D. Xiong, *Applications of light emitting diodes in health care,* Ann. Biomed. Eng. **45**, 2509 (2017).

[22] J. Dong and G. Zhang, *Identification and robust control of the nonlinear photoelectrothermal dynamics of LED systems,* IEEE Trans. Ind. Electron. **64**, 2215 (2017).

[23] D. Zorov, M. Juhaszova, and S. Sollott, *Mitochondrial reactive oxygen species (ROS) and ROS-induced ROS release,* Physiol. Rev. **94**, 909 (2014).

[24] F. Cieplik, A. Pummer, J. Regensburger, K. Hiller, A. Späth, L. Tabenski, W. Buchalla, and T. Maisch, *The impact of absorbed photons on antimicrobial photodynamic efficacy,* Front. Microbiol. **6**, 706 (2015).

[25] R. Buchanan, M. Golden, and R. Whiting, *Differentiation of the effects of ph and lactic or acetic acid concentration on the kinetics of listeria monocytogenes inactivation,* J. Food Prot. **56**, 474 (1993).

[26] P. Hansen, V. Pereyra, and G. Scherer, *Least Squares Data Fitting with Applications* (Johns Hopkins University Press, 2012).

[27] M. Hamblin, *Mechanisms and mitochondrial redox signaling in photobiomodulation,* Photochem. Photobiol. **94**, 199 (2018).

[28] A. Batlle, *Porphyrins, porphyrias, cancer and photodynamic therapy – a model for carcinogenesis,* J. Photochem. Photobiol., B **20**, 5 (1993).

[29] R. Sampaio, M. Silva, A. Batista, and N. B. Neto, *Investigation of the photophysical and eletrochemical properties of a free base tetrapyridyl porphyrin with meso carbon linked ruthenium (II) groups,* J. Photochem. Photobiol., A **315**, 98 (2016).

[30] L. Brovko, A. Meyer, A. Tiwana, W. Chen, H. Liu, C. Filipe, and M. Griffiths, *Photodynamic treatment: A novel method for sanitation of food handling and food processing surfaces,* J. Food Prot. **72**, 1020 (2009).

[31] T. Leong, R. Vierstra, and W. Briggs, *A blue light-sensitive cytochrome-flavin complex from corn coleoptiles. further characterization,* Photochem. Photobiol. **34**, 697 (1981).

4

ANALYZING EFFICACY AND SAFETY OF ANTI-FUNGAL BLUE LIGHT THERAPY VIA KERNEL-BASED MODELING THE REACTIVE OXYGEN SPECIES INDUCED BY LIGHT

Objective: *The goal of this study is to investigate the efficacy, safety, and mechanism of ABL for inactivating* Candida albicans *(*C. albicans*), and to determine the best wavelength for treating candida infected disease, by experimental measurements and dynamic modeling.* Methods: *The changes in reactive oxygen species (ROS) in* C. albicans *and human host cells under the irradiation of 385, 405, and 415nm wavelengths light with irradiance of* $50mW/cm^2$ *were measured. Moreover, a kernel-based nonlinear dynamic model, i.e., nonlinear autoregressive with exogenous inputs (NARX), was developed and applied to predict the concentration of light-induced ROS, whose kernels were selected by a newly developed algorithm based on particle swarm optimization (PSO).* Results: *The ROS concentration was increased respectively about 10-12 times in* C. albicans *and about 3-6 times in human epithelial cells by the ABL treatment with the same fluence of* $90J/cm^2$*. The NARX models were respectively fitted to the data from the experiments on both types of cells. Besides, four different kernel functions, including Gaussian, Laplace, linear and polynomial kernels, were compared in their fitting accuracies. The errors with the Laplace kernel turned out to be only 0.2704 and 0.0593, as respectively fitted to the experimental data of the* C. albicans *and human host cells.* Conclusion: *The results demonstrated the effectiveness of the NARX modeling approach, and revealed that the 415nm light was more effective as an anti-fungal treatment with less damage to the host cells than the 405 or 385nm light.* Significance: *The kernel-based NARX model identification algorithm offers opportunities for determining the effective and safe light dosages in treating various*

fungal infection diseases.

4.1. INTRODUCTION

Candida albicans (*C. albicans*) is widely found in nature and commonly occurs as a superficial infection on mucous membranes, such as the mouth, throat, gut, and vagina [1–3], and is the most common fungal pathogen of humans [4]. *Candida* species are naturally found in 10%-20% of women [5]; and 75% of women have at least one episode of vulvovaginal candidiasis (VVC) [6]. About 90% of the overall cases of VVC are caused by *C. albicans* [7]. Topical antimycotic drugs and more convenient oral azole agents are the main treatments for VVC [2]. However, *C. albicans* has shown increased resistance to these drugs [8, 9]. Therefore it is crucial to identify new ways to treat fungal infections. Photodynamic therapy (PDT) has been investigated as an alternative to treat localized infectious diseases due to the rapid action and avoidance of drug resistance by these pathogens [10]. Similar to PDT, anti-fungal blue light (ABL) therapy relies only on endogenous photosensitizers (PS) of the pathogens, and are hence safer to use.

The hypothesized mechanism of the antimicrobial effect of PDT and ABL is that light photons excite either exogenous PS in the former case or endogenous PS in the latter, which, in turn, produces highly toxic ROS in cells [11]. Due to the type and content of endogenous PS in different cells, their susceptibility to ABL can also be different. Fungal PS content is normally higher than that of human cells. Thus, fungi are more susceptible to ABL and, hence, ABL has been widely investigated for treating fungal infections. For instance, the inactivation rate of *C. albicans* by ABL was 42-fold faster than human keratinocytes [12]. A dynamic model was developed based on the viability of *C. albicans* and vaginal epithelial (V. E.) cells during ABL irradiation, which demonstrated that the shorter ABL wavelength around 410nm achieved a higher anti-fungal effect than 450nm [13]. The safety of ABL in treating candidal vaginitis was investigated in [14], which found that the blue light at 405nm preferentially induced more death to the pathogenic cells than to the human V.E. cells. Furthermore, no genotoxicity of blue light to the V. E. cells was observed at the dosage for inactivating the pathogen. However, to the best of our knowledge, no study has focused on the ROS concentrations in fungi and host cells for treating VVC, i.e., *C. albicans* and the V. E. cells.

Some first-principle models have been proposed based on the PDT mechanism. For instance, modeling of the dynamic changes in ROS concentrations has been well investigated [15–17], which are highly nonlinear models. These first-principle models precisely fit the dynamic changes in ROS concentrations. However, a dynamic model for the ROS concentrations to treat fungal infections has not been established. This can be attributed to the difficulty of using a first principle model, i.e., the parameters are related to the PS characteristics, which are determined experimentally. For instance, up to 21 parameters are required to describe the process related to the PS [16]. In contrast, although ABL is believed to be caused by the PS that naturally exist in fungal cells, whose types and amounts are usually unknown, first-principle modeling becomes even more challenging than modeling PDT.

Generally, to handle the challenges in modeling by first principles, system identification methods (SIM) that estimate models from the data measured from complex dynamic processes have been well developed in control theory literature [18]. In the recent

decades, the applications of SIM methods to identify biological and biomedical systems, which are usually highly nonlinear, coupled and chaotic, have also been witnessed, e.g., in modeling arterial windkessel [19], aortic pressure [20] and electrodermal activity [21]. To deal with the nonlinearities in biological systems, various structured nonlinear dynamic model identification methods have been investigated, e.g., Wiener model [20] and NARMAX model [22].

Despite all the aforementioned efforts in biomedical system identification, identifying a nonlinear dynamic model for anti-fungal blue light therapy has not yet been targeted in the literature. However, some attempts have been made to handle similar problems. For instance, a closed-loop control scheme has been implemented to track the photobleaching trajectory during PDTs in [23], which is an ON/OFF controller designed without any model. A data-driven modeling method has been investigated in [24], which basically approximates the nonlinear PDT dynamics by a linear integrator model. Until now, no nonlinear model based on experimental data has been established. Furthermore, no study has analyzed the effects and safety of ABL therapy based on a dynamic ROS model.

In this study, ROS levels were measured in *C. albicans* and V. E. cells under irradiation from three LED light sources of different wavelengths, including 385, 405, and 415nm. Then, these experimental data were fitted using a popular kernel-based method [25], e.g., nonlinear autoregressive with exogenous inputs (NARX) modeling [26, 27]. The advantage of using kernel based learning is mainly the treatment of the nonlinearity of a complex dynamic process by linearly combining a set of kernels. Furthermore, sparse kernel modeling can be applied to select the best kernel centers from the training samples. One popular approach is based on random selection by minimizing some cost functions, e.g., using the repeating weighted boosting search (RWBS) algorithm [28], which is an evolutionary algorithm based on weight boosting search. In this method, the kernel parameters and the centers are chosen by minimizing a MSE objective function. However, to search one regressor, it needs to be repeated for multiple times with initial random sampling, before finally converges to the global optimum. This iteration may reduce the algorithm efficiency. To avoid this iteration and hence improve the efficiency, we choose the particle swarm optimization (PSO) algorithm instead of RWBS, which is a proven method for its fast searching speed [29]. The fitting accuracy of the NARX model to ROS detected in *C. albicans* and V. E. cells was satisfactory.

The contributions of the current study are three-fold. First, we conducted blue light stimulation experiments and measured the induced ROS in both *C. albicans* and V. E. cells, in response to three different wavelengths (385, 405, and 415nm). Second, a kernel-based NARX model was developed and applied to the experimental data. Moreover, a new PSO-based kernel selection algorithm was proposed and applied to improve this NARX model. The fitting accuracy demonstrated the effectiveness of the modeling approach. Third, we analyzed the optimal ABL wavelength for treating VVC using the simulation results.

4.2. METHODS

4.2.1. NARX MODELING

The "kernel trick" was used to reduce the experimental burden and treat the nonlinearity of the dynamics of ROS accumulation, e.g., using the NARX model with a suitable kernel function. A general discrete-time nonlinear system is described as:

$$\boldsymbol{y}_k = f(\boldsymbol{y}_{k-1}, \cdots, \boldsymbol{y}_{k-n_y}, \boldsymbol{u}_{k-1}, \cdots, \boldsymbol{u}_{k-n_u}) + \boldsymbol{\epsilon}_k, \tag{4.1}$$

where $\boldsymbol{u}_k \in \mathbb{R}^m, \boldsymbol{y}_k \in \mathbb{R}^\ell$, and $\boldsymbol{\epsilon}_k \in \mathbb{R}^\ell$ are respectively the input, output, and noise vector at time instant k; $f(\cdot)$ is a nonlinear function; and $n_y, n_u \in \mathbb{N}$ represent respectively the output and input delays, e.g., $\boldsymbol{u}_{k-n_u} = \boldsymbol{u}_k \cdot z^{-n_u}$, with z^{-1} standing for the one step delay operator. The noise ϵ_k is zero-mean white Gaussian with a covariance matrix Σ, i.e., $\epsilon_k \sim \mathcal{N}(0, \Sigma)$.

To model the changes in intracellular ROS concentrations by Eq. (4.1), let $\boldsymbol{y}_k$ denote the ROS concentration at the kth time sampling point, f denote the ROS generation process, $\hat{\boldsymbol{y}}_k$ denote the estimated value of $\boldsymbol{y}_k$, and u_k denote the irradiance of the light. In ABL, the irradiance is usually kept constant during the entire treatment process, i.e., $u_k \equiv u, \forall k > 0$. So in (4.1), it is not necessary to consider the inputs at different delay steps. Instead, the effect of the exogenous input u can be considered as a step response. That is, when the light is switched on, the excitation of the ROS starts. Mathematically, this process can be rewritten as:

$$\boldsymbol{y}_k = \hat{\boldsymbol{y}}_k + \boldsymbol{\epsilon}_k = \tilde{f}_u(\boldsymbol{y}_{k-1}, \cdots, \boldsymbol{y}_{k-n_y}) + \boldsymbol{\epsilon}_k, \tag{4.2}$$

where $\tilde{f}_u(.)$ is defined as:

$$\tilde{f}_u(\boldsymbol{y}_{k-1}, \cdots, \boldsymbol{y}_{k-n_y}) = \begin{cases} 0, & \text{if } u_k = 0 \\ f(\boldsymbol{y}_{k-1}, \cdots, \boldsymbol{y}_{k-n_y}), & \text{if } u_k = u \end{cases}. \tag{4.3}$$

For simplicity, collect the sequence of $\boldsymbol{y}_{k-1}, ..., \boldsymbol{y}_{k-n_y}$ into a column vector, and denote it as $\boldsymbol{x}_k = [\boldsymbol{y}_{k-1}^T, ..., \boldsymbol{y}_{k-n_y}^T]^T$. Let the number of kernel basis functions be n. Eq. (4.2) was changed into the following regression model using some suitable functions that approximate $f(\cdot)$ with arbitrary accuracy,

$$\hat{\boldsymbol{y}}_k = \Sigma_{i=1}^n \boldsymbol{w}_i g_i(\boldsymbol{x}_k) \tag{4.4}$$

where $\boldsymbol{w}_i \in \mathbb{R}^\ell, i = 1, \cdots, n$ are the corresponding weighting vectors; and $g_i, \in \mathbb{R}, i = 1, \cdots, n$, is a kernel function chosen from the popular candidates listed in Tab. 4.1.

The number of I/O data samples was denoted by N. The output vectors were collected into a matrix as:

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 & \boldsymbol{y}_2 & \boldsymbol{y}_3 & \cdots & \boldsymbol{y}_N \end{bmatrix}^T \in \mathbb{R}^{N \times \ell} \tag{4.5}$$

Similarly, the outputs of the kernel functions excited by $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N$ were collected into the following regressor matrix

Table 4.1: Some popular kernel functions, with $\boldsymbol{x}_i$ being the center and a_i, b_i, and c_i being the parameters of the i-th kernel.

Kernel	$g_i(\boldsymbol{x})$
Linear	$a_i \boldsymbol{x}^T \boldsymbol{x}_i + b_i$
polynomial	$(a_i \boldsymbol{x}^T \boldsymbol{x}_i + b_i)^{c_i}, a_i > 0, c_i \in \mathbb{N}^+$
Gaussian(RBF)	$exp\left(-\frac{\|\boldsymbol{x}-\boldsymbol{x}_i\|_2^2}{2a_i^2}\right)$
Laplace	$exp\left(-a_i \|\boldsymbol{x}-\boldsymbol{x}_i\|\right)$

$$\begin{aligned} \boldsymbol{G} &= \begin{bmatrix} g_1(\boldsymbol{x}_1) & \cdots & g_n(\boldsymbol{x}_1) \\ \vdots & & \vdots \\ g_1(\boldsymbol{x}_N) & \cdots & g_n(\boldsymbol{x}_N) \end{bmatrix} \in \mathbb{R}^{N\times n} \\ &= \begin{bmatrix} \boldsymbol{g}_1 & \boldsymbol{g}_2 & \boldsymbol{g}_3 & \cdots & \boldsymbol{g}_n \end{bmatrix}, \end{aligned} \tag{4.6}$$

where $\boldsymbol{g}_i = [g_i(\boldsymbol{x}_1), \cdots, g_i(\boldsymbol{x}_N)]^T$.

With the aforementioned definition, Eq. (4.4) can be written in compact form as

$$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{w}, \tag{4.7}$$

where $\boldsymbol{w} = \begin{bmatrix} \boldsymbol{w}_1 & \boldsymbol{w}_2 & \boldsymbol{w}_3 & \cdots & \boldsymbol{w}_n \end{bmatrix}^T \in \mathbb{R}^{n\times \ell}$.

With a given regressor matrix, the only parameter in Eq. (4.7) to be estimated is the weight matrix $\boldsymbol{w}$. On the other hand, the regressor matrix is determined by the kernels with the set of parameters $\{\boldsymbol{x}_i, a_i, b_i, c_i, i = 1, \cdots, N\}$. The method to determine these kernels will be detailed later in what follows.

First, to solve $\boldsymbol{w}$, we use the QR factorization of $\boldsymbol{G}$, i.e.,

$$\boldsymbol{G} = \boldsymbol{P}\boldsymbol{A}, \tag{4.8}$$

where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is an upper diagonal matrix; $\boldsymbol{P} \in \mathbb{R}^{N\times n}$ is as orthogonal matrix. Therefore, Eq. (4.7) can be rewritten as

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{G}\cdot\boldsymbol{w} \\ &= \boldsymbol{P}\boldsymbol{A}\cdot\boldsymbol{w} \\ &= \boldsymbol{P}\cdot\boldsymbol{\theta} \end{aligned} \tag{4.9}$$

where $\boldsymbol{\theta} = \boldsymbol{A}\boldsymbol{w} \in \mathbb{R}^{n\times\ell}$, and can simply be solved as $\boldsymbol{\theta} = \boldsymbol{P}^T\boldsymbol{y}$.

Learning a NARX model from data requires estimating the weights $\boldsymbol{\theta}$ and the parameters of the kernel functions, e.g., the kernel center $\boldsymbol{x}_i$. In this study, a kernel selection method based on particle swarm optimization (PSO) was applied to select a subset of the best kernels from the full set of candidate kernels. This selection algorithm is detailed in the Appendix. A.1. Incorporated by this selection algorithm, the entire NARX model identification algorithm is summarized in Appendix A.2.[1]

[1]The codes implementing these algorithms are available from https://drive.google.com/drive/folders/1tSwPW5aBiS3KYAWcTKwRapLasZ2ZqdSD?usp=sharing, or by contacting with the author.

4.2.2. LED LIGHT SOURCE

Three different types of LEDs with the specified peak wavelengths respectively at 385nm, 405nm, and 415nm were applied in this work. The types of these LEDs and their main parameters are listed in Tab. 4.2. Their spectral power density (SPD) curves, as measured by a Maya2000Pro spectrometer (Ocean Optics, US), are depicted in Fig. 4.1(a). In this figure, every SPD curve is normalized with respect to its integral over the range of the measured wavelength, i.e., with each normalized SPD curve integrating to 1.

The LED light sources were designed following the procedures in [30], and were driven by a constant current source with PWM current level control to stabilize the output irradiance [31]. The LED chips are arranged in a 1.5cm-by-1.5cm square as a 4-by-4 array. The three LED light sources can deliver an irradiance of $50 mW/cm^2$ uniformly within a $60 cm$-diameter circle. Fig. 4.1(c) shows the simulated irradiance distribution, where the average irradiance in the 6cm-diameter circle is $49.39 mW/cm^2$, with a relative variation of only 6.56%. The irradiance was measured and confirmed by a PM100D power meter with a S120VC probe (Thorlabs, US). The schematic diagram of the electrical drive and control system and the experimental setup are illustrated in Figs. 4.1(b) and 4.1(d).

Table 4.2: The LEDs used and their main parameters, where FWHM represents full width at half maximum.

type	wavelength	FWHM
Vishay VLMU3500-385-120	385nm	10nm
Kingbright ATDS3534UV405B	405nm	15nm
LUMILEDS LUXEON LHUV-0415-A070	415nm	13.7mn

4.2.3. CULTURE CONDITIONS FOR *C. albicans* AND THE V. E. CELLS

The human host cell line used in this study was the vaginal epithelial cell strain (VK2/E6E7 ATCC CRL-2616, ATCC, Manassas, VA, USA). The cell line was incubated in Dulbecco's modified Eagle's medium (Gibco, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum at 37^oC in a humidified atmosphere with 5% CO_2.

The *C. albicans* used in this study was the 3147 (IFO 1594) strain (ATCC). The fungal strain was cultured in tryptic soy broth at 26^oC. To maintain the concentration of the fungal suspension within the same range, the absorption of the suspension was measured at 550nm using the U-3900H spectrophotometer (Hitachi, Tokyo, Japan) before all experiments. The absorption levels measured by this equipment were always controlled in the range of 2.3-2.5, which corresponded to a fungi density of 10^7 CFU/ml.

4.2.4. ROS ASSAY

The fungi were centrifuged, separated from the medium, and dissolved in a 1,000-fold dilution of the ROS fluorescent probe (DCFH-DA assay kit, Beyotime Institute of Biotechnology, Beijing, China) in phosphate buffered saline. After incubating the suspension at 37^oC in a shaker for 30 min, the suspension was centrifuged three times to remove the redundant probe. Then, the suspension was seeded into a 96-well plate and was irradiated with one of the three LED light sources.

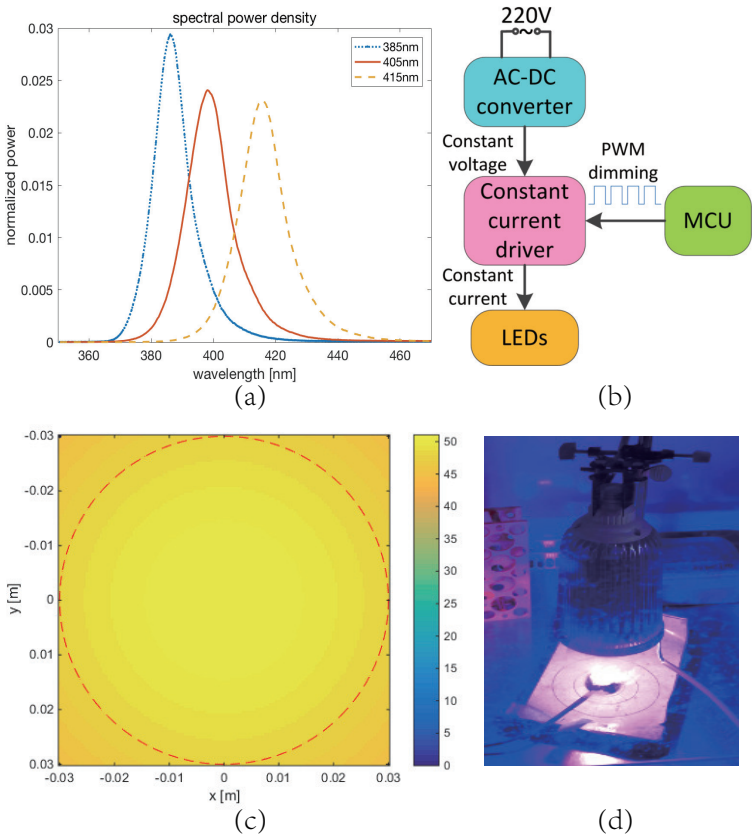


Figure 4.1: LED light source design: (a) normalized SPD of the four types of LEDs, (b) electrical scheme, (c) simulated irradiance distribution, (d) photo of the experimental setup.

During the experiment, the cell suspension in one well was resuspended and taken out of the plate consecutively at 0, 5, 10, 15, 20, 25, and 30 min. The removed suspension in one well was placed in the dark; while the rest wells were kept under the irradiation of the light. Therefore, the fluence received by each well was 0, 15, 30, 45, 60, 75, and 90 J/cm^2, respectively, considering the light irradiance of 50 mW/cm^2. The intracellular ROS level of the cell suspension was measured immediately after the light treatment using a VL0L0TD0 Varioskan LUX microplate reader (Thermo Fisher, Waltham, MA, USA), with excitation and emitting wavelengths of 488 and 525 nm, respectively.

Finally, a sequence of ROS fluorescent levels up to 30 min were obtained. The entire experimental method was repeated three times independently. In total, triplicate experiments were performed. In each repeating experiment, a new cell strain was thawed and incubated, then seeded into a new 96-well plate, and finally irradiated following the aforementioned procedures.

4.2.5. MODELING THE VIABILITY OF *C. albicans* AND V. E. CELLS

Although the light-induced ROS causes cytotoxicity, the susceptibility to the ROS of the two cells can still be different, in the sense that the same amount of ROS may cause different viability reduction in both types of cells. To further investigate this issue, the survival rates of *C. albicans* and V. E. cells affected by ABL were modeled.

The viability models take the following form, which is a piecewise function [13, 32], including a shoulder at the beginning of the ABL to represent the accumulating process of the light-induced cytotoxic ROS.

$$N_{c,e}(t) = \begin{cases} N_{c,e}(0), & t < \tau \\ N_{c,e}(0) \cdot e^{-\kappa(t-\tau)}, & t \geq \tau \end{cases}, \tag{4.10}$$

where $N_{c,e}(t)$ is the survival rate at time t; the subscripts "c, e" respectively represent the *C. albicans* and V. E. cells; κ is the decaying rate coefficient; and τ is the time constant of when the inactivation starts.

4.2.6. STATISTICS

In the ROS assays, the raw data were processed to produce the mean and standard deviation for each treatment time interval. The significance of ROS levels and viability of cells were tested by the Student's t-test. The values of $P < 0.05$ were considered statistically significant.

4.3. RESULTS

4.3.1. MEASUREMENTS OF THE ROS CONCENTRATIONS IN *C. albicans* AND V. E. CELLS DURING ABL IRRADIATION

The time sequences of the ROS concentrations in the *C. albicans* and V. E. cells were measured using all three light sources. The measured fluorescence levels were processed as $y = \frac{R}{R_0}$, where R_0 is the initial ROS fluorescence of the cells, i.e., the ROS level not altered by light, as the control group; R is the measured ROS fluorescence level after the ABL irradiation, as the treatment group; and y is the processed relative fluorescence level. We denoted the relative change in the *C. albicans* ROS level as y_c; and similarly denoted that of V. E. cells as y_e. The y_c and y_e values irradiated by ABL of the three wavelengths from 0 to 30 min are plotted in Fig. 4.2.

After *C. albicans* was exposed to light for 5 min, the relative ROS concentration y_c increased significantly in all cases (P<0.05); y_c increased 10-12 times within 30 min. The y_e value of the V. E. cells increased significantly after 5 min of light exposure in all cases (P<0.05); y_e was increased 3-6 times within 30 min. Considering the experimental data from the both cells together, y_c was significantly higher than y_e after 15 min of irradiation in all cases (P<0.05).

4.3.2. NARX MODELING AND COMPARISON BETWEEN PSO AND RWBS KERNEL SELECTION METHODS

The NARX was used to predict the dynamic changes in *C. albicans* ROS concentrations; and the data from the experiments with different wavelengths were fitted. The relative

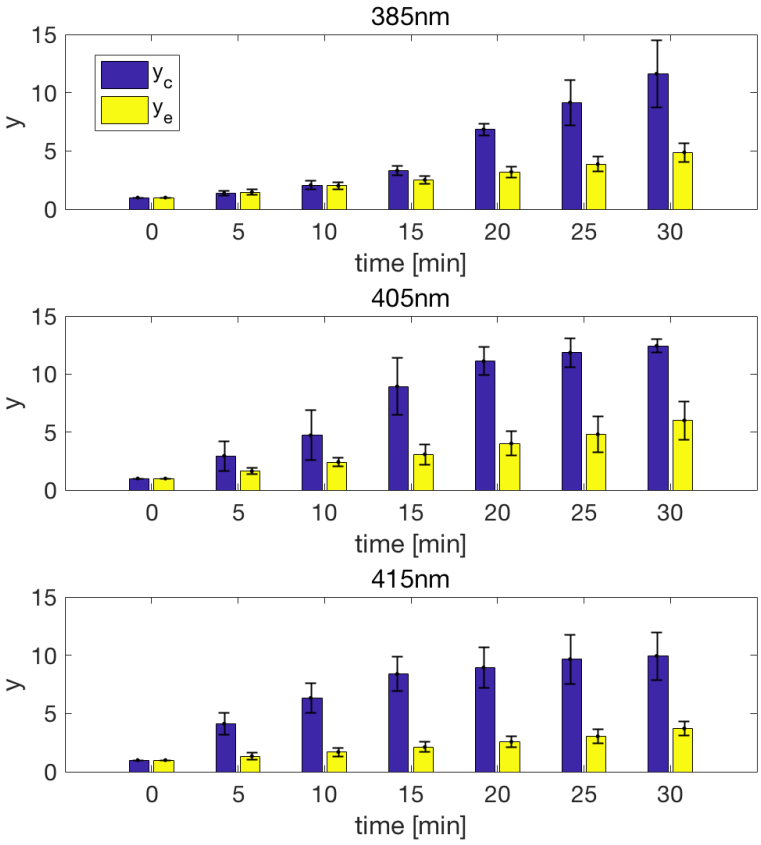


Figure 4.2: Rrelative ROS levels of *C. albicans* and V. E. cells under irradiation with three wavelengths of ABL

ROS concentration of *C. albicans* at the kth time sampling point was denoted by $y_c(k)$. Here, k is from 1 to 7; and corresponded *C. albicans* was irradiated for 0 to 30 min with a step size of 5 min. Let $n_y = 2$, i.e., in a second order nonlinear dynamic form. Then, the NARX model can be written as:

$$y_c(k) = \tilde{f}_u(y_c(k-2), y_c(k-1)) \tag{4.11}$$

The ROS concentrations of the cells irradiated by 385, 405, and 415nm wavelength light were measured as time-series data. These time-series data were used to estimate the NARX model. More specifically, seven data sampling points were processed with a delay of two steps to generate five pairs of I/O data as required in Eq. (4.11).

The NARX model was estimated as described in Sec. 4.2.1 using these five pairs of I/O data. Then, the first two data sampling points were taken to predict the third point. After that, the iteration was continued by plugging the predicted points into the right hand side of Eq. (4.11).

Here, we compared the kernel selection methods respectively by the PSO and RWBS

algorithm [28], on fitting the NARX model to the data from the ABL experiment on the *C. albicans* under the irradiation of the 415nm light. The parameters of RWBS algorithm were chosen the same as those used in [28], i.e., population size $Q = 40$, outer loop repeat times $N = 7$, and inner loop repeat times $K = 600$. The PSO parameters were $Q = 40$, $N = 1$ and $K = 5$, which can yield comparable model accuracy as the RWBS algorithm, as listed in Tab. 4.3. The weighting factors of Eq. (A.1) in Appendix A.1 were set to $d_1 = 0.6, d_2 = 0.6, d_3 = 0.5$ empirically.

For this problem, the average time cost of running the RWBS algorithm was about 2.7 seconds; while the average time cost of the PSO algorithm was only about 0.8 seconds. Both of them ended up with 2 best kernels.

Table 4.3: Comparison of the efficacy and accuracy of different kernel selection algorithms of NARX model.

Selection algorithm	RMSE	Time cost (s)
RWBS	0.5642	2.747
PSO	0.5407	0.8364

4.3.3. NARX MODELING BASED ON *C. albicans* ROS

To select a best kernel for $g_i(x)$ from the candidates listed in Tab. 4.1, the NARX models were compared in simulations. They were compared in terms of the fitting error of the RMSEs of NARX models, which are shown in Tab. 4.4.

Table 4.4: Comparison of the accuracy of the NARX models with different kernels fitted to the data from the experiments with different wavelengths.

Wavelength	RMSE with different kernels			
[nm]	**RBF**	**Laplace**	**Linear**	**Polynomial**
385	0.8676	0.2019	0.8662	0.6051
405	0.7319	0.3163	0.9407	0.5748
415	0.2108	0.2931	0.2393	0.2394
Average	0.6035	0.2704	0.6821	0.4731

Since the Laplace kernel performs the best in the simulation, we chose the NARX model with the Laplace kernel to model the process in the following. This predicting results of the NARX model are shown in Fig. 4.3.

4.3.4. NARX MODELING BASED ON MEASURED ROS IN V. E. CELLS

Similar to section III.C, the NARX model was used to predict the dynamic changes in the ROS concentrations of V. E. cells. The NARX model took the form of Eq. (4.11). As the increase in y_e was similar to that of y_c, we chose NARX with the Laplace kernel to model the process. The modeling accuracy of the NARX models against the data collected from different wavelengths is listed in Tab. 4.5 and plotted in Fig. 4.4.

Table 4.5: The RMSEs of the NARX model with the Laplace kernel estimated by the experimental data from V. E. cells

Wavelength	385nm	405nm	415nm	**Average**
RMSE	0.0544	0.0827	0.0409	0.0593

4.3.5. ANALYZING THE SAFETY AND EFFICACY OF ABL THERAPIES VIA NARX MODELS

In the above sections, NARX models were established to estimate the ROS concentrations induced by single-wavelength ABL for *C. albicans* and V. E. cells. In this section, the ROS concentrations in the cells were compared using different NARX models. More specifically, we plotted the tuple of (y_c, y_e) stimulated by the same fluence ranging from

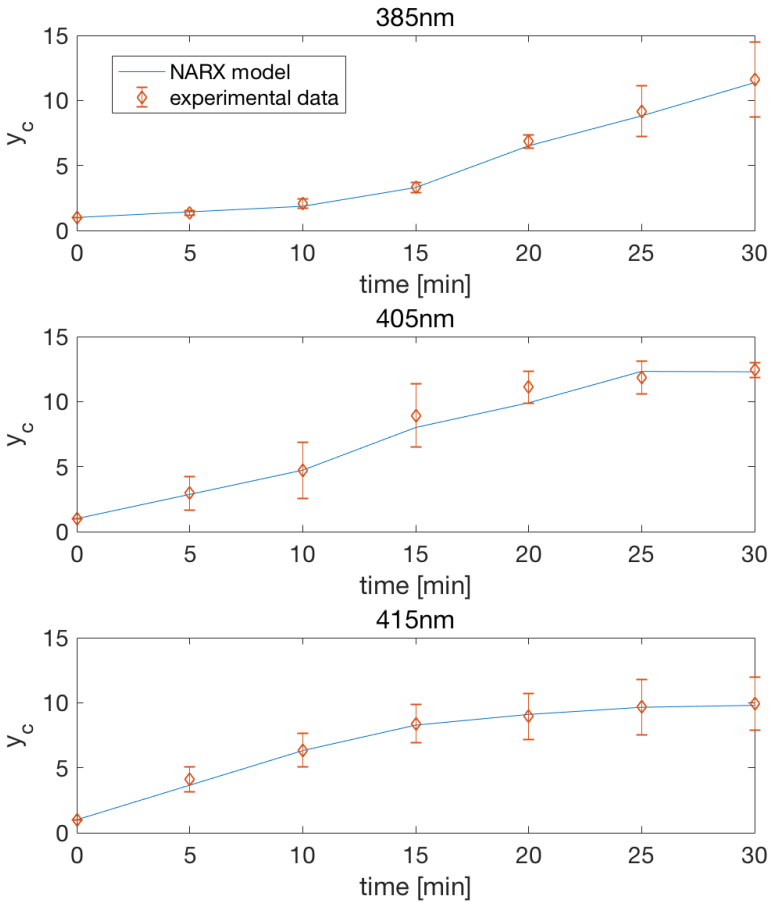


Figure 4.3: The NARX model predicting the measured relative ROS concentrations in *C. albicans* from the 385, 405 and 415nm light irradiation experiment, respectively.

0 to 90 J/cm^2 with a step size of 15 J/cm^2. Here, fluence was calculated as the product of irradiance and irradiation time. The results are shown in Fig. 4.5. A dash line $y_c = y_e$ was plotted to better reflect the trends in the figure. It can be clearly observed that the ROS generation was greater in the *C. albicans* than in the V. E. cells for each wavelength and at a same amount of fluence. As shown in Fig. 4.5, the red dash curve (415nm) is the lowest one among the three curves, which indicates the consistently lowest ROS generation in the V. E. cells under the 415nm light exposure. Therefore, the 415nm light had less damage to the human host cells than the 385nm and 405nm. In comparison, the 405nm light was the most harmful to the V. E. cells.

4.3.6. COMPARISON OF THE NARX MODELS WITH LINEAR AR MODELS IN FITTING THE EXPERIMENTAL DATA

To further show the advantage of the nonlinear NARX modeling approach, linear autoregressive (AR) models (see [18]) were also identified to predict the dynamic changes of ROS in the *C. albicans* and V. E. cells. Similar to the NARX model, the delay step was also set to 2. The RMSEs of the AR models are listed in Tab. 4.6. Clearly, the accuracy of the AR models was much lower than that of the NARX models with the Laplace kernels, which have been reported in Tab. 4.4 and 4.5, respectively for *C. albicans* and V. E. cells.

Table 4.6: The RMSE of the AR model fitted to data of all the three wavelengths.

Wavelength [nm]	**C. A.**	**V. E.**
385	3.6702	0.9402
405	3.2734	1.6248
415	2.1904	1.3077
Average	3.0477	1.2909

4.3.7. GENERALIZED NARX MODEL FOR SHORT WAVELENGTH ABL THERAPY

In this section, we developed a generalized NARX model to fit the experimental data from all three wavelengths. Here, the NARX model identified by the experimental data with one wavelength was used to fit the experimental data to all three wavelengths. The average fitting RMSEs of the NARX to the *C. albicans* experimental data are listed in Tab. 4.7. The NARX with a polynomial kernel estimated by 385 and 405nm data performed well with a good generalized capability. Similarly, the polynomial kernel also showed good generalized capability for the V. E. cells data. The NARX model estimated by the 405nm wavelength data performed the best of all.

4.3.8. THE SURVIVAL RATES OF *C. albicans* AND V. E. CELLS AFFECTED BY ABL

By the aforementioned observations from Fig. 4.5, the 415nm was the best choice for treating *C. albicans* infected VVC; while the 405nm was the worst. Therefore, we only need to take these two extreme cases to study the inactivating effects on both *C. albicans* and V. E. cells. It shall be emphasized that since the main contribution of this work is to

develop and verify the kernel-based NARX models of light-induced ROS concentrations, extensively repeating the well-studied viability assays is out of its scope. On the other hand, in our previous work [13], such experiments with exactly the same types of cells and experimental settings were already performed, which had resulted in the viability data, i.e., $N_c(t)$ and $N_e(t)$ for $t = 0, 5, 15, 20, 25$ min, from respectively the 405nm and 415nm light irradiation. The details of the applied materials and methods can be found therein.

Eq. (4.10) was fitted to these viability data, and applied to estimate the lethal doses of the V. E. cells. The fitted parameters are listed in Tab. 4.8, with which the lethal doses of the V. E. cells were finally estimated, and are listed in Tab. 4.9. Note that calculating LD90 of the V. E. cells is especially relevant, because at the max fluence applied to the *C. albicans*, i.e., after the irradiation by either of the wavelengths for 25min, N_c reached the order of 10^{-3}. However, at the same dosage, the survival rates of the V. E. cells were still a bit higher than 0.5, by either the 405nm or 415nm light. Since the treatment target had already been achieved, it was not necessary to further apply higher fluence to the V. E. cells. Therefore, the LD90 calculated by the model is a reasonable estimate of the dosage for killing up to 90% V. E. cells.

Furthermore, the tuples of (N_c, N_e) caused by the same dosages are illustrated in Fig. 4.6. Besides, linear lines were fitted to show the descending trends of the V. E. cells, as the viability rates of the *C. albicans* decreased. The slopes of the linear lines were -0.4733 and -0.4031, as respectively fitted to the experimental data of the 405nm and 415nm.

4.4. DISCUSSION

4.4.1. EXPERIMENTAL AND SIMULATION RESULTS

In this study, the ROS generated during anti-fungal light irradiation experiments were measured with three different wavelengths. The main objectives were to compare the effects of the three different wavelengths on inducing ROS, and to build a NARX model to predict the changes in ROS levels.

In our experiments, light-induced ROS increased significantly in *C. albicans*, demonstrating the effectiveness of the ABL therapy. On the other hand, the ROS in the V. E. cells was also increased during blue light irradiation. However, the increased ROS ratio of V. E. cells was much lower than that of *C. albicans*, which demonstrates the safety of ABL therapy.

The ROS level produced by the 405nm light was slightly higher than that of the other two wavelengths, indicating a higher effectiveness for inducing ROS in *C. albicans*. This

Table 4.7: The accuracy of the estimated NARX models fitted to data of all the three wavelengths. The experimental data used were the *C. albicans* time-series ROS concentrations.

Wavelength	**Average RMSE with different kernels**			
[nm]	**RBF**	**Laplace**	**Linear**	**Polynomial**
385	1.3707	1.1629	2.1573	0.9875
405	1.8357	2.8710	2.0509	0.9402
415	1.5546	1.1629	1.6395	1.5785

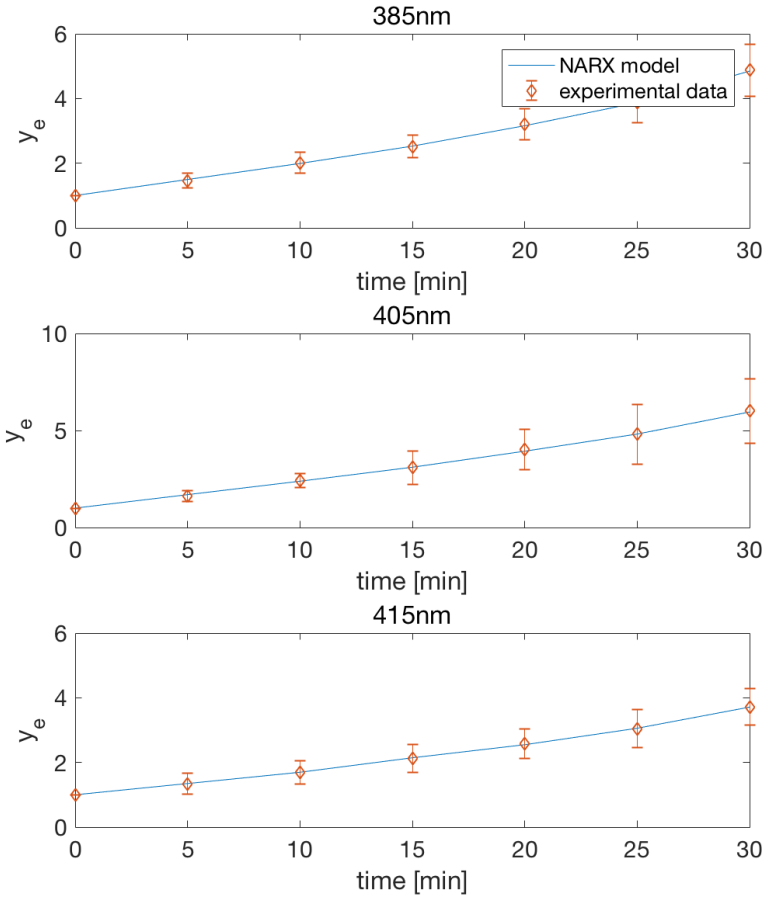


Figure 4.4: The NARX model predicting the measured relative ROS concentrations in V. E. cells from the 385, 405 and 415nm light irradiation experiment, respectively.

coincided with the finding that the main PS in *C. albicans* is porphyrins [12, 33], whose peak absorption wavelength is 405nm. Thus, the 405nm blue light excited the porphyrins more effectively than the other two wavelengths.

V. E. cells were also most affected by the 405nm blue light. Although the main endogenous PS in V. E. cells is flavin adenine dinucleotide, whose peak absorption wavelength is about 450nm [34], they also include considerable coproporphyrin content [14]. Besides, the molar extinction coefficient of coproporphyrin is about 45 times higher than that of flavin adenine dinucleotide [35, 36]. Thus, 405nm light also induced the ROS effectively in V. E. cells.

The main difficulty with ROS assay experiments is that the initial cellular ROS level is difficult to control. The initial cellular ROS level is proportional to the cell concentration, which is strictly controlled by absorption of the fungal suspension. Also, the culture conditions, e.g., media and passage or culture time, can affect the amount of cellular

Table 4.8: Fitted parameters of the viability model (4.10).

	405nm		415nm	
	τ	κ	τ	κ
C. A.	769	85.9	450	3.80e-3
V. E.	56	224.3	526	5.81e-4

Table 4.9: Estimated lethal doses, LD50 and LD90, respectively for 50% and 90% viability reduction.

	405nm	415nm
LD50 $[J/cm^2]$	87.1	85.9
LD90 $[J/cm^2]$	282.5	224.3

PS [37]; therefore, affecting the initial ROS level and ability of ROS to accumulate. In this study, the culture conditions were strictly controlled. All experiments were performed with fresh-thawed fungal strains, which were cultured for the same durations.

On the other hand, the performance of the polynomial kernel was best among the four kernels, when the generalized NARX model was used to fit the experimental data of all three wavelengths. The fitting accuracy of the linear kernel was not satisfactory, which was attributed to the nonlinearity of the ROS accumulation process. For instance, some kinetic models of light-induced ROS have been investigated [38–40] and all of these kinetic models demonstrated high nonlinearity.

In Sec. 4.3.7, a generalized NARX model was developed to predict the ROS induced by the three wavelengths of light. In fact, the 385, 405 and 415nm wavelengths include almost all of the short wavelength blue light. Thus, this generalized NARX model was useful to predict the ROS concentration with any short wavelength blue light. We did not perform experiments with long wavelength ABL, since it does not inactivate fungal strains [13, 34]. Thus, short wavelength ABL is a potential and effective treatment; and our generalized NARX model has provided a satisfactory prediction of the light-induced ROS concentration in responses to short wavelength blue light.

4.4.2. COMPARISON OF THE EFFECTS BASED ON THE EXPERIMENTAL AND SIMULATION RESULTS OF ALL THREE WAVELENGTHS

The ROS concentrations in the different cells were compared in Sec. 4.3.5. The dash line in Fig. 4.5 represents $y_c = y_e$, indicating the same ROS concentrations in *C. albicans* and V. E. cells in response to the same fluence of light. All three curves were always below it, which explains why the relative increase of ROS in V. E. cells was lower than that of *C. albicans*. More specifically, by the same relative ROS concentrations in *C. albicans* (y_c=8), the increase in the ROS ratio in V. E. cells was 2.6, 3, and 2.13, respectively with the 385, 405, and 415nm light. Obviously, the 415nm light has a more effective anti-fungal function with less damage to the human host cells than 385 and 405nm.

However, the same amount of ROS may still cause different viability reduction in different types of cells. The slopes of the fitted lines in Fig. 4.6 help to further investigate this issue, which turned out to be −0.4733 and −0.4031, as respectively fitted to the experimental data of the 405nm and 415nm reported in our previous work [13]. The slope

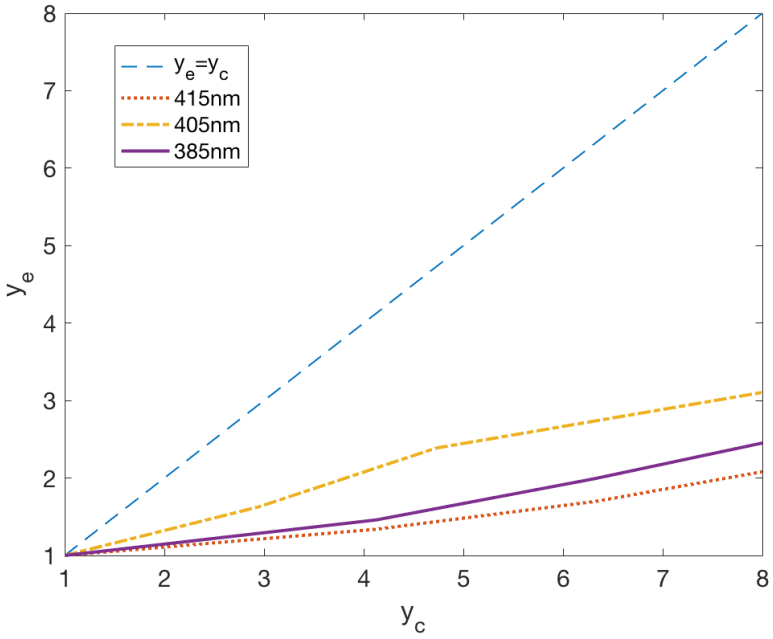


Figure 4.5: Comparing the light-induced ROS concentration using the NARX models identified from the data of two types of cells

reflects the resistance of the V. E. cells to the light relatively to that of the *C. albicans*. The lower its absolute value, the more resistant the host cells to the light than the pathogens. Therefore, Fig. 4.6 shows that the 415nm light has a more effective anti-fungal function with less damage to the human host cells than the 405nm. Putting the observations from both Fig. 4.5 and Fig. 4.6 together, it can be concluded that the 415nm is the best wavelength to treat this disease, in terms of both the efficacy and safety.

Notably, the 405nm ABL was the most effective wavelength to induce the ROS in *C. albicans* and V. E. cells. However, the 415nm is the better wavelength for treating VVC. This result reminds us that we should not only focus on the ROS accumulation by pathogens, but pay attention to the damage to host cells. The peak absorption of specific pathogens, such as *C. albicans*, is 405nm, determined by the endogenous PSs. However, this is not to say that the 405nm is always the best ABL wavelength to treat candida infections.

4.4.3. THE SAFETY OF ABL

Clinically applicable anti-fungal strategies shall selectively inactivate pathogenic fungi, while sparing the normal host cells and tissues. As a safety study of ABL, both the ROS accumulation and survival rates were investigated and compared in *C. albicans* and V. E. cells. The increased ROS concentration in the V. E. cells was much lower than that in the *C. albicans*, implying the favorable selectivity of the ABL. During the 415nm light exposure, no significant inhibition of the V. E. cells was observed within the beginning 10min; while 47% of the *C. albicans* was killed at that time. In the end, over 58% of the V.

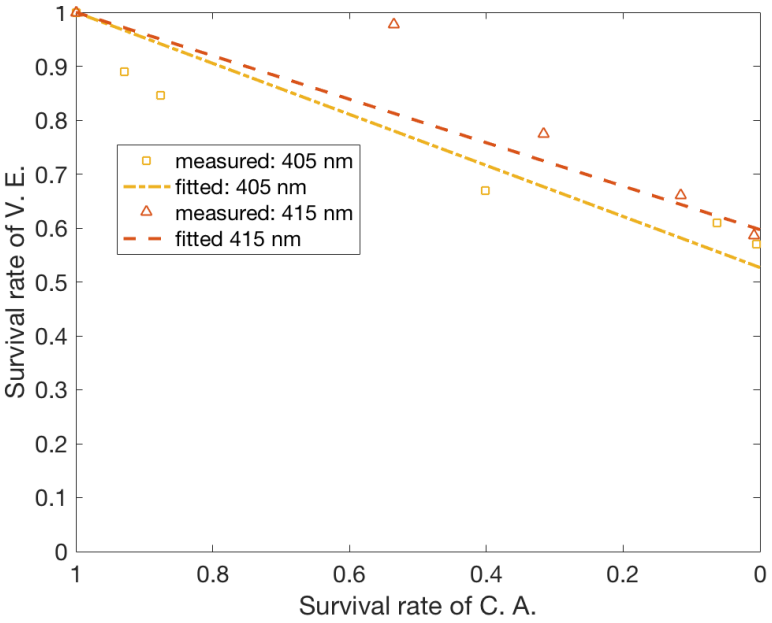


Figure 4.6: The survival rates for the *C. albicans* and V. E. cell in the irradiation of three wavelengths, i.e., 385, 405, and 415nm.

E. cells survived after the 415nm light irradiation, which almost completely inactivated the *C. albicans*. The aforementioned results suggest that there exists a therapeutic window, where the *C. albicans* can be selectively inactivated, while the majority of V. E. cells are preserved.

Furthermore, the maximum ABL fluence applied in this study was $90 J/cm^2$. Such a fluence is a safe dose without causing any genotoxicity to human host cells. It has been reported by other authors that no ABL induced DNA damage occurred in the epithelial cells up to the $216 J/cm^2$ fluence of 405nm light [14]. For the more sensitive human retinal pigment epithelial (RPE) cells, it has also been observed that the induced damage to the mitochondrial DNA in the RPE cells was less at the fluence of $60 J/cm^2$ than that at $30 J/cm^2$ [41]. This may be attributed to the DNA repair system that was activated after the $30 J/cm^2$ light treatment. In fact, it was proved that the mitochondria are capable of repairing oxidative DNA damage to some extent, e.g., the damage to bases and single-strand breaks [42].

4.5. CONCLUSION

In this study, we measured the changes in intracellular ROS of *C. albicans* and V. E. cells for 30 min with irradiation of 385, 405, and 415nm light. Furthermore, we proposed a modeling scheme using a kernel-based NARX structure, whose kernels were selected by a newly developed algorithm based on PSO optimization. This NARX model was used to fit the experimental data. High fitting accuracy was achieved by the model, demonstrat-

ing the effectiveness of the proposed modeling technique. Both the experimental data and the numerical results from the NARX model indicated that the ROS ratio of V. E. cells was always lower than that of *C. albicans,* demonstrating the safety of the ABL therapy. Moreover, a key conclusion was that the 415nm wavelength blue light was the most effective wavelength, with the least damage to V. E. cells. More importantly, the proposed kernel-based NARX model identification algorithm can also be applied to determine the effective and safe light dosages in treating other types of fungal infection diseases.

REFERENCES

[1] T. Arendorf and D. Walker, *The prevalence and intra-oral distribution of candida albicans in man,* Arch. Oral Biol. **25**, 1 (1980).

[2] J. D. Sobel, *Vaginitis,* N. Engl. J. Med. **337**, 1896 (1997).

[3] A. Erdogan and S. S. Rao, *Small intestinal fungal overgrowth,* Curr. Gastroenterol. Rep. **17**, 16 (2015).

[4] R. A. Calderone and W. A. Fonzi, *Virulence factors of candida albicans,* Trends Microbiol. **9**, 327 (2001).

[5] J. D. Sobel, *Candidal vulvovaginitis,* Clin. Obstet. Gynecol. **36**, 153 (1993).

[6] R. Hurley and J. de Louvois, *Candida vaginitis,* Postgrad. Med. J. **55**, 645 (1979).

[7] F. C. Odds *et al.*, *Candida and candidosis: a review and bibliography* (Bailliere Tindall, 1988).

[8] M. A. Pfaller, *Antifungal drug resistance: mechanisms, epidemiology, and consequences for treatment,* Am. J. Med. **125**, S3 (2012).

[9] P. Vandeputte *et al.*, *Antifungal resistance and new strategies to control fungal infections,* Int. J. Microbiol. **2012** (2012).

[10] T. Maisch, *A new strategy to destroy antibiotic resistant microorganisms: antimicrobial photodynamic treatment,* Mini Rev. Med. Chem. **9**, 974 (2009).

[11] R. Yin *et al.*, *Light based anti-infectives: ultraviolet C irradiation, photodynamic therapy, blue light, and beyond,* Curr. Opin. Pharmacol. **13**, 731 (2013).

[12] Y. Zhang *et al.*, *Antimicrobial blue light inactivation of Candida albicans: In vitro and in vivo studies,* Virulence **7**, 536 (2016).

[13] T. Wang *et al.*, *Blue light therapy to treat candida vaginitis with comparisons of three wavelengths: an in vitro study,* Lasers Med. Sci. **35**, 1329 (2020).

[14] Y. Wang *et al.*, *Photoinactivation of neisseria gonorrhoeae: a paradigm-changing approach for combating antibiotic-resistant gonococcal infection,* J. Infect.Dis. **220**, 873 (2019).

[15] K. Wang *et al.*, *A comprehensive mathematical model of microscopic dose deposition in photodynamic therapy,* Med. Phys. **34**, 282 (2007).

[16] M. Kim *et al.*, *On the in-vivo photochemical rate parameters for pdt reactive oxygen species modeling,* Phys. Med. Biol. **62**, R1 (2017).

[17] T. Zhu *et al.*, *Macroscopic modeling of the singlet oxygen production during pdt,* in *Optical Methods for Tumor Treatment and Detection: Mechanisms and Techniques in Photodynamic Therapy XVI*, Vol. 6427 (International Society for Optics and Photonics, 2007) p. 642708.

[18] L. Ljung, *System Identification - Theory for the User* (Prentice-Hall, 1987).

[19] T. Kind *et al.*, *Estimation of three- and four-element windkessel parameters using subspace model identification,* IEEE. Trans. Biomed. Eng. **57**, 1531 (2010).

[20] A. Patel *et al.*, *Aortic pressure estimation using blind identification approach on single input multiple output nonlinear wiener systems,* IEEE. Trans. Biomed. Eng. **65**, 1193 (2018).

[21] M. R. Amin and R. T. Faghih, *Sparse deconvolution of electrodermal activity via continuous-time system identification,* IEEE. Trans. Biomed. Eng. **66**, 2585 (2019).

[22] Y. Gu *et al.*, *Nonlinear modeling of cortical responses to mechanical wrist perturbations using the NARMAX method,* IEEE. Trans. Biomed. Eng. **In Press** (2020).

[23] J. Tylcz *et al.*, *Realtime tracking of the photobleaching trajectory during photodynamic therapy,* IEEE. Trans. Biomed. Eng. **64**, 1742 (2017).

[24] J. Dong and T. Wang, *Data driven modeling of the reactive oxygen species stimulated by photon energy in light therapies,* IEEE Access **8**, 18196 (2020).

[25] S. Chen *et al.*, *Orthogonal least squares learning algorithm for radial basis function networks,* IEEE Trans. Neural Netw. Learn. Syst. **2**, 302 (1991).

[26] J. Dong, *Recursive NARX model identification of nonlinear chemical processes with matrix invertibility analysis,* Comput. Chem. Eng. **121**, 574 (2019).

[27] W. Lee *et al.*, *NARX modeling for real-time optimization of air and gas compression systems in chemical processes,* Comput. Chem. Eng. **115**, 262– (2018).

[28] S. Chen *et al.*, *NARX-based nonlinear system identification using orthogonal least squares basis hunting,* IEEE Trans. Contr. Syst. Technol. **16**, 78 (2007).

[29] C. Ou and W. Lin, *Comparison between PSO and GA for parameters optimization of pid controller,* in *2006 International conference on mechatronics and automation* (IEEE, 2006) pp. 2471–2475.

[30] J. Dong and Z. Zhang, *Design of LED light for stimulating cells in the study of light therapies,* in *Proc. 15th SSLChina: IFWS* (IEEE, 2018) pp. 1–4.

[31] J. Dong and D. Xiong, *Applications of light emitting diodes in health care,* Ann. Biomed. Eng. **45**, 2509 (2017).

[32] R. Buchanan *et al.*, *Differentiation of the effects of pH and lactic or acetic acid concentration on the kinetics of listeria monocytogenes inactivation,* J. Food Protect. **56**, 474 (1993).

[33] V. Plavskii *et al.*, *Porphyrins and flavins as endogenous acceptors of optical radiation of blue spectral region determining photoinactivation of microbial cells,* J. Photochem. Photobiol. B, Biol. **183**, 172 (2018).

[34] Y. Wang *et al.*, *Antimicrobial blue light inactivation of neisseria gonorrhoeae: Roles of wavelength, endogenous photosensitizer, oxygen, and reactive oxygen species,* Lasers Surg. Med. **51**, 815 (2019).

[35] C. Rimington, *Spectral-absorption coefficients of some porphyrins in the soret-band region,* Biochem. J. **75**, 620 (1960).

[36] J. A. Lewis and J. C. Escalante-Semerena, *The FAD-dependent tricarballylate dehydrogenase (TcuA) enzyme of salmonella enterica converts tricarballylate into cis-aconitate,* J. Bacteriol. **188**, 5479 (2006).

[37] J. Fyrestam *et al.*, *Influence of culture conditions on porphyrin production in aggregatibacter actinomycetemcomitans and porphyromonas gingivalis,* Photodiagnosis Photodyn. Ther. **17**, 115 (2017).

[38] N. Lopez *et al.*, *Tumor reactive ringlet oxygen approach for Monte Carlo modeling of photodynamic therapy dosimetry,* J. Photochem. Photobiol. B, Biol. **160**, 383 (2016).

[39] G. Kareliotis *et al.*, *Assessment of singlet oxygen dosimetry concepts in photodynamic therapy through computational modeling,* Photodiagnosis Photodyn. Ther. **21**, 224 (2018).

[40] K. Wang, *Explicit dosimetry for photodynamic therapy: macroscopic singlet oxygen modeling,* J. Biophotonics **3**, 304 (2010).

[41] B. Godley *et al.*, *Blue light induces mitochondrial dna damage and free radical production in epithelial cells,* J. Biol. Chem. **280**, 21061 (2005).

[42] V. A. Bohr and R. M. Anson, *Mitochondrial dna repair pathways,* J. Bioenerg. Biomembr. **31**, 391 (1999).

5

MULTIPLE MODEL BASED NARX MODELING OF EGFR SIGNALING NETWORKS IN MAMMALIAN CELLS

The epidermal growth factor receptor (EGFR) is an important factor that regulates cell growth, survival, proliferation and differentiation. The dynamics of the concentrations of the proteins in an EGFR signaling network are nonlinear and chaotic. Modeling by first principles results in Michaelis-Menten type equations. These equations involve many unknown parameters, e.g. reaction rates and initial conditions, which are difficult to be measured or estimated. To avoid these modeling difficulties, we propose an approach using kernel-based NARX (nonlinear autoregressive with exogenous inputs) techniques. The novelty of this work is two-fold. First, we propose a method based on particle swarm optimization (PSO) to select optimal kernels of the NARX model. Second, we propose an adaptive estimation approach based on the multiple model principle to handle the chaotic nature of the EGFR network. The convergence of the model adaptation is also analyzed. Simulation results have verified the effectiveness of the proposed methods in modeling the nonlinear and chaotic EGFR signaling network.

5.1. INTRODUCTION

The epidermal growth factor receptor (EGFR) belongs to the family of protein-tyrosine kinase receptors. EGFR is widely distributed on the surface of mammalian epithelial cells, fibroblasts, glial cells and keratinocytes[1]. EGFR signaling pathway plays an important role in cell growth, proliferation and differentiation. In [2, 3], the first principle dynamic models of the EGFR signaling network are proposed, which are highly nonlinear and chaotic.

To describe the dynamics of chemical or biological processes, time-series models are widely applied [4–7]. Like the EGFR signaling network, these process are nonlinear, and contain time-varying parameters due to aging, exogenous disturbances. Various nonlin-

ear time-series models have been studied to handle these difficulties. In [4, 7, 8], NARX models are parameterized to identify the nonlinear chemical processes therein. On the other hand, to tackle the nonlinearities in complex dynamical processes, kernel and orthogonal basis function based learning [9–11] and system identification [12] techniques are also becoming more and more popular.

In kernel based methods, sparse kernel modeling has been widely applied in selecting the best kernel centers from the training samples. A popular method for kernel selecting is the orthogonal least squares (OLS) algorithm [13, 14]. A limitation of this method is choosing the centers with a fixed common Gaussian RBF kernel variance. However, this kernel variance is critical for the generalization of the model. Another approach is selecting the kernel centers and common variance at the same time by global optimizing methods. For example, in [15], a genetic algorithm (GA) has been used to tune the parameters of the kernel though optimizing the model performance.

Like GA, PSO is also a global optimizer. The PSO algorithm is originally attributed to [16, 17], and has been used in computer science and chemical engineering [18–20]. The major advantage of PSO is the fast convergence speed. In [21] the convergence speed between GA and PSO is compared, which shows that PSO converges faster and performs better than GA does. Another global optimizer called repeating weighted boosting search (RWBS) [22] has also been applied in kernel selection [23]. In this method, the kernel variance and the center are chosen by minimizing an MSE objective function. However, to search one regressor, it needs to repeated for multiple times with initial random sampling before finally converges to the global optimum. This iteration may reduces the algorithm efficiency. To avoid this iteration and hence improve the efficiency, we propose to use PSO algorithm instead of RWBS for kernel selection. The PSO algorithm has been used to optimize the parameters, such as input or output delays of NARX models [24, 25]. However, PSO has not yet been applied in selecting kernels in NARX models in this literature. In this work, we propose to select the centers and the variance of the kernels by PSO algorithm, and will hence call it "NARX-PSO" for brevity.

On the other hand, the time-series data of the concentrations in the EGFR signaling network shows a chaotic characteristic, because the key parameters (such as the initial concentration of proteins) vary with time. In adaptive control, one of the popular method is to use multiple models (MM) based adaptive structure [26–29], where the local models are convexly combined. The weighting coefficients are also updated during the model adaptation which can be regarded as the second level adaptation, and can dramatically increase the predicting accuracy.

In this paper, the time-series data are generated by simulating the Michaelis-Menten type first principle model of the concentrations of various proteins in the EGFR signaling network. From the data, we identify NARX models that can predict the dynamic changes of a protein or the phosphorylated fraction of a protein group therein. Since the initial concentration of the EGF can largely change the dynamic trajectories of any other proteins in this network, this process exhibits a chaotic nature. We will therefore build up multiple NARX models corresponding to different initial conditions, and adapt the weight of each model to track the trajectories initiated at any arbitrary EGF concentration.

This paper is organized as follows. Section 2 introduces the first principle model of

the EGFR signaling network. In Section 3, NARX model and kernel selection method are proposed. Section 4 is devoted to developing the adaptive multiple model approach. Section 5 gives the simulation results to show the effectiveness of the proposed approach. Section 6 concludes the paper.

5.2. EGFR SIGNALING NETWORK

The epidermal growth factor receptor (EGFR) [1] belongs to the family of protein-tyrosine kinase receptors, which regulates cell growth, survival, proliferation, and differentiation. Because of EGF binding, the Grb2 (growth factor receptor-bound protein 2) starts binding and being activated, Shc (generic shell script compiler) and PLCγ (Phospholipase C, gamma) are also phosphorylated.

The EGFR signaling network is illustrated in Fig. 5.1. While EGF binds to the extracellular domain, it causes a series of reactions including 23 proteins or receptors. For brevity, we only introduce the beginning part, i.e. the reactions from EGF to RP, to describe this network. This subnetwork starts with EGF binding to the extracellular domain of the monomeric EGFR (designated as R) in step 1, i.e. the reaction 1 in Fig. 5.1, and then forms the EGF-EGFR complex (designated as Ra). EGF binding drives the association of the two receptor monomers into an activated receptor dimer (step 2). A 2:2 Ra complex is the predominant form of the receptor dimer (designated as R2). The phosphorylation of tyrosine residues by receptor tyrosine kinase is combined into a single step 3, yielding a form designated as RP [2]. Then, the reactions with RP and other groups are activated. A quantitative description of the short term EGFR signaling of these process is described in [2], which contains a set of kinetic models describing the temporal changes of the concentration of all the ingredients. The reacting rate in step n is denoted as v_n. The dynamics of each step are described by the rate equations and kinetic models are provided in the Appendix. A.3.

In this work. we treat this first principle model as the underlying true process, and simulate it to generate the time-series data to identify the NARX model. Note that only the step 1 process reacts in the extracellular domain. The concentration of EGF is hence adjustable. Different initial concentrations of EGF can be used to simulated the first principle model, and to test the chaotic nature of the dynamics.

5.3. NARX MODELING AND KERNEL SELECTION

5.3.1. PRELIMINARIES

Clearly from the model described in the Appendix. A.3, the dynamics of the EGFR signaling network is nonlinear and highly coupled. This complicated nonlinearity can be handled by the "kernel trick", e.g. by the NARX model with RBF kernels. For brevity, the preliminaries refer to NARX is omitted, due to the same NARX model was already introduced in Chapter 4.

5.3.2. KERNEL SELECTION BY PSO

The objective of kernel selection is to select a subset of n_s ($n_s \ll n$) best kernels from the full set of n candidate kernels. One popular approach to choose the centers is based on random selection by minimizing some cost functions, e.g. using the RWBS algorithm in

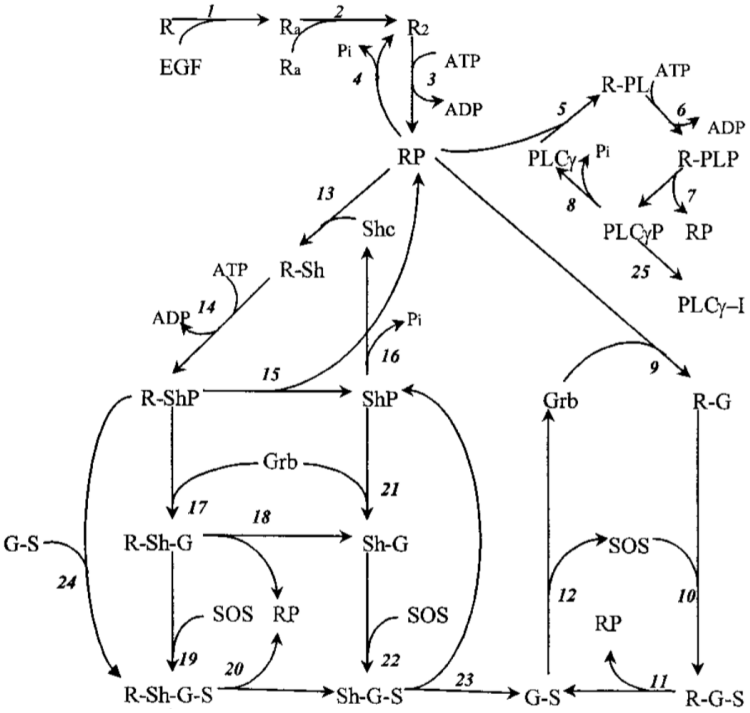


Figure 5.1: The EGFR signaling network. Containing 25 reaction steps and 23 proteins.

[23], which is an evolutionary algorithm based on weight boosting search. In [22], it has been shown to have a similar convergence speed as GA. Here, we propose to use the PSO algorithm, another evolutionary algorithm that mimics the movement of the organisms in a bird flock, in the NARX kernel selection, and will show its advantage in terms of the speed over RWBS later in the simulation studies.

For simplicity of notations and also for the specific case of the EGFR signaling network studied in this work, we will especially consider the single output case, i.e. $\ell = 1$ and hence $\boldsymbol{y} \in \mathbb{R}^N$, $\boldsymbol{\theta} \in \mathbb{R}^n$. However, it shall be mentioned that the proposed method is not restricted to single output case. For the NARX kernel selection, define a cost function as J_t, where t stands for the number of selected regressors. The initial cost is denoted as $J_0 = \boldsymbol{y}^T \boldsymbol{y}$. According to Eq. (4.9) one can write

$$J_0 = \boldsymbol{\theta}^T \boldsymbol{P}^T \boldsymbol{P} \boldsymbol{\theta} = \Sigma_{i=1}^{n} \boldsymbol{p}_i^T \boldsymbol{p}_i \theta_i^2.$$

The idea is to search all the columns in $\boldsymbol{P}$, and find the one that reduces the cost value most, if being removed from J_t.

$$J_t = J_{i-1} - \boldsymbol{p}_i^T \boldsymbol{p}_i \theta_i^2, \tag{5.1}$$

where $\boldsymbol{p}_i$ is chosen from $\boldsymbol{P}$. This procedure can be terminated if

$$J_T < \xi \tag{5.2}$$

is satisfied, where ξ is a chosen positive scalar; and the subscript "T" represents that T regressors are chosen.

The PSO algorithm is proposed to choose the orthogonal columns in Eq. (5.1). In PSO, the decision variables are regarded as particles. These particles move around in the search space. The movements of the particles are guided by their own best known position in the search space as well as the best known position of the entire swarm.

First, all particles are dispersed uniformly. The movement, also known as velocity, is denoted as V, whose initial value can be randomly chosen. Let $\Psi_j, V_j \in \mathbb{R}^S$ denotes the j-th particle and its corresponding velocity, where S denotes the dimension of Ψ_j and V_j. The movements are adapted by the following formula [16, 17].

$$V_j = wV_j + c_1(\Psi^* - \Psi_j) + c_2(\Psi_j^* - \Psi_j), \tag{5.3}$$

where w, c_1 and c_2 are the weighting factors; Ψ^* and Ψ_j^* are the global best particle and local best particle respectively. The value of V_j shall be restricted into the so-called "setting region" as follows

$$V_{j,s} = \begin{cases} V_{min,s}, & \text{if } V_{j,s} \le V_{min,s}, s = 1,2,\cdots,S \\ V_{max,s}, & \text{if } V_{j,s} \ge V_{min,s}, s = 1,2,\cdots,S \\ V_{j,s}, & \text{else} \end{cases}, \tag{5.4}$$

where the subscript "s" denotes the s-th element of V_j, V_{min}, V_{max}.

After each movement, the position of the j-th particle is updated by

$$\Psi_{j+1} = \Psi_j + V_j, \tag{5.5}$$

which is then restricted into the search space as follows

$$\Psi_{j,s} = \begin{cases} \Psi_{min,s}, & \text{if } \Psi_{j,s} \le \Psi_{min,s}, s = 1,2,\cdots,S \\ \Psi_{max,s}, & \text{if } \Psi_{j,s} \ge \Psi_{min,s}, s = 1,2,\cdots,S \\ \Psi_{j,s}, & \text{else} \end{cases}, \tag{5.6}$$

where the subscript "s" denotes the s-th element of $\Psi_j, \Psi_{min}, \Psi_{max}$.

The search will stop when the cost function of Ψ^* is satisfied or the maximum generation is reached.

In the specific EGFR signaling model, the I/O data are time-series data of the concentrations of proteins. For each protein, the time-series data of the concentration is a vector. To model each protein, the output dimension is $\ell = 1$, and hence y_k, θ_i take the scalar form. The parameters in the NARX model to be optimized are the kernel center $\boldsymbol{x}_i$ and the kernel variance σ_i. Thus, for kernel selection, the particle shall include the index of the center and its corresponding variance; i.e. $\Psi_j = [i_j, \sigma_{i_j}]^T$, where i_j stands for the index of the kernel center that is included in the j-th particle. Once these kernel parameters are fixed, the corresponding $\boldsymbol{p}_i$ and θ_i can be calculated by the standard Gram-Schmidt procedure; and the cost function can be calculated by Eq. (5.1).

Algorithm 1 (NARX Kernel Selection by PSO)

Outer loop *(search the t-th regressor g_t, $t = 1, 2, \cdots, T$)*
Initialize the value of population size denoted as Q; the weighting factors w, c_1 and c_2; the searching space Ψ_{min}, Ψ_{max}; and the movement setting region V_{min}, V_{max}.
for*(t=1; t ≤ T; t=t+1)*

1. *Initialize the population randomly in the searching space and randomly generate an initial velocity matrix V inside the setting region.*

2. ***for****(j=1; j ≤ Q; j=j+1)*

   (a) *Choose the kernel index i_j from the particle $\mathbf{\Psi}_j = [i_j, \sigma_{i_j}]^T$. Compute the regressor vector $\boldsymbol{g}_{i_j} = [g_{i_j}(\boldsymbol{x_1}), \cdots, g_{i_j}(\boldsymbol{x_N})]^T$ by Tab. 4.1 and Eq. (4.6), as a candidate while selecting the t-th regressor. Then orthogonalize it by the standard Gram-Schmidt procedure:*

   (b) *if* $t = 1$

$$\begin{aligned} \boldsymbol{u}_j &= \boldsymbol{g}_{i_j} \\ \boldsymbol{p}_{i_j} &= \frac{\boldsymbol{u}_j}{\|\boldsymbol{u}_j\|_2} \end{aligned}$$

   (c) *else*

$$\begin{aligned} \boldsymbol{u}_j &= \boldsymbol{g}_{i_j} - \Sigma_{r=1}^{t-1}(\boldsymbol{g}_{i_j}^T \cdot \boldsymbol{p}_r) \cdot \boldsymbol{p}_r \\ \boldsymbol{p}_{i_j} &= \frac{\boldsymbol{u}_j}{\|\boldsymbol{u}_j\|_2} \end{aligned} \tag{5.7}$$

   (d) *Calculate the corresponding θ_{i_j} and the cost function in terms of $\boldsymbol{p}_{i_j}, \theta_{i_j}$ by*

$$\theta_{i_j} = \frac{\boldsymbol{p}_{i_j}{}^T y}{\boldsymbol{p}_{i_j}{}^T \boldsymbol{p}_{i_j}} \tag{5.8}$$

$$J_t^j = J_t^{j-1} - \boldsymbol{p}_{i_j}{}^T \boldsymbol{p}_{i_j} \theta_{i_j}{}^2 \tag{5.9}$$

   where the superscript j denotes the the j-th particle. i.g. J_t^j is the cost after choose $\mathbf{\Psi}_j = [i_j, \sigma_{i_j}]^T$ as the t-th regressor.

   end for

3. *Find the best particle, as the one corresponding to the minimal J_t^j, i.e.*

$$j^* = \min_j \{J_t^1, J_t^2, \cdots, J_t^Q\}.$$

   And set the global best as $\Psi^* = \Psi_{j^*}$.

4. *Initialize the local best particles as*

$$\Psi_j^* = \Psi_j, j = 1, 2, \cdots, Q$$

The PSO inner loop:

for*(k=1; k ≤ T; k=k+1)*

for*(j = 1; j ≤ Q, j = j + 1)*

(a) Calculate V_j by Eq. (5.3) and restrict its value into the setting region by Eq. (5.4).

(b) Update Ψ_j by Eq. (5.5) and restrict its value into the search space by Eq. (5.6).

(c) Calculate the cost function value J_t^j as in Eq. (5.9) for the new particle.
Update the global best particle Ψ^* and the local best particles Ψ_i^* as follows.
If Ψ_j is better than Ψ^*, set $\Psi^* = \Psi_j$.
If Ψ_j is better than Ψ_j^*, set $\Psi_j^* = \Psi_j$.

end for

Stop the inner loop when the maximum epoch is reached or $J_t^{j^*} < \xi_P$.

end for

End of inner loop

The best particle find in this epoch is Ψ^*. Set $J_t = J_t^{j^*}$, $\boldsymbol{g}_t = \boldsymbol{g}_{i_{j^*}}$, $\boldsymbol{p}_t = \boldsymbol{p}_{i_{j^*}}$ and $\boldsymbol{\theta}_t = \boldsymbol{\theta}_{i_{j^*}}$.
Remove the selected index j^* from the candidates of the regressor indices.

end for

End of outer loop

5.4. MULTIPLE MODEL AT DIFFERENT INITIAL CONDITIONS AND WEIGHT ADAPTATION

According to the experimental study in [2], the initial concentrations of EGF can be different. Because of this difference, the concentrations of the other proteins show chaotic dynamics. In this section, we will handle this chaos by a multiple-model approach. Let h be the number of the local models, and $\boldsymbol{\alpha}$ be the weighting coefficient, i.e.

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_h \end{bmatrix}^T \in \mathbb{R}^h$$

The final output is convexly combined by the results from all h local models.

$$\hat{\boldsymbol{y}}(k) = \sum_{i=1}^{h} \alpha_i \hat{\boldsymbol{y}}_i(k) \tag{5.10}$$

To this end, first, train all the local models individually by different datasets. Here, a dataset is a set of I/O data corresponding to a certain initial concentration of EGF. Second, update $\tilde{\boldsymbol{\alpha}}$ by the following rule [30].

$$\tilde{\boldsymbol{\alpha}}(k) = \tilde{\boldsymbol{\alpha}}(k-1) - \tilde{\boldsymbol{E}}\tilde{\boldsymbol{E}}^T\tilde{\boldsymbol{\alpha}}(k-1) + \tilde{\boldsymbol{E}}e_h(k), \tag{5.11}$$

where $\boldsymbol{e}_i$, $\tilde{\boldsymbol{\alpha}}$, $\tilde{\boldsymbol{E}}$ are defined as follows.

$$\begin{aligned} \boldsymbol{e}_i &= \boldsymbol{y} - \hat{\boldsymbol{y}}_i & (5.12)\\ \tilde{\boldsymbol{\alpha}} &= \begin{bmatrix}\alpha_1 & \alpha_2 & \ldots \alpha_{h-1}\end{bmatrix} & (5.13)\\ \tilde{\boldsymbol{E}} &= \begin{bmatrix}\boldsymbol{e}_1 - \boldsymbol{e}_h & \boldsymbol{e}_2 - \boldsymbol{e}_h & \ldots \boldsymbol{e}_{h-1} - \boldsymbol{e}_h\end{bmatrix} & (5.14)\end{aligned}$$

According to Eq. (5.11), the convergence of the updating rule is ensured by the following condition

$$\tilde{\boldsymbol{E}}\tilde{\boldsymbol{E}}^T < 2,$$

Obviously, this condition depends on the error vector $\tilde{\boldsymbol{E}}$, which can take arbitrary values depending on the underlying system dynamics. Hence, we add a scaling factor β to $\tilde{\boldsymbol{E}}\tilde{\boldsymbol{E}}^T$ in order to enforce this condition. To make the updating rule still fulfill the equality of the error equations, i.e. $\tilde{e}(k) = \tilde{\boldsymbol{E}}\tilde{\boldsymbol{\alpha}}(k)$ as originally defined in [26], scaling by β is performed as follows

$$\tilde{\boldsymbol{\alpha}}(k) = \tilde{\boldsymbol{\alpha}}(k-1) - \beta\left[\tilde{\boldsymbol{E}}\tilde{\boldsymbol{E}}^T\tilde{\boldsymbol{\alpha}}(k-1) + \tilde{\boldsymbol{E}}e_h(k)\right], \quad (5.15)$$

Now, the condition for the convergence of the updating rule is reduced to

$$\beta \cdot \tilde{\boldsymbol{E}}\tilde{\boldsymbol{E}}^T < 2. \quad (5.16)$$

In this structure, h NARX models denoted by $M_i, i = 1, \cdots, h$ are convexly combined by h convex coefficients, each taking values within the range [0,1], and $\sum_{i=1}^{h} \alpha_i = 1$.

For simplicity and also for the specific case of the EGFR network, let us consider the case of two local models. Each of the local model is a NARX model identified from the data simulated with one initial EGF concentration. The local model M_1 is trained with the minimum EGF concentration; while M_2 is trained at its maximum. In this case, $y(k), e(k), \tilde{E}(k)$ become scalars. Eqs. (5.10)-(5.14) can hence be rewritten as

$$\begin{aligned} \hat{y}(k) &= \alpha_1(k)\hat{y}_1(k) + \alpha_2(k)\hat{y}_2(k) & (5.17)\\ e(k) &= y(k) - \hat{y}(k)\\ e_1(k) &= y(k) - \hat{y}_1(k)\\ e_2(k) &= y(k) - \hat{y}_2(k)\\ \tilde{E}(k) &= e_1(k) - e_2(k)\\ &= \hat{y}_2(k) - \hat{y}_1(k)\end{aligned}$$

Here, $\hat{y}(k)$ is the estimate of the final output; $e_i(k)$ is the error for the i-th model at time k; and $\hat{y}_i(k)$ is the output of the i-th local model.

According to the condition of (5.16), in this case, the condition for the convergence of the updating rule is reduced to

$$\beta\tilde{E}^2 < 2. \quad (5.18)$$

This means that the difference between the two local models is bounded by $\pm\sqrt{2}/\beta$. In practice, on can select a proper value of β to ensure that Eq. (5.18) holds; and α converges.

We will now show that if α_1 converges, i.e. if $\lim_{k\to\infty}\alpha_1(k) = \lim_{k\to\infty}\alpha_1(k-1)$, then $y(k) = \hat{y}(k)$ holds. For brevity, we will omit the "lim" operator in the following derivation. According to Eq. (5.15), as $k \to \infty$ one can write

$$\beta\tilde{E}^2\alpha_1(k-1) = -\beta\tilde{E}e_2(k).$$

Therefore, one has

$$\alpha_1(k) = \alpha_1(k-1) = -\frac{e_2(k)}{\tilde{E}} = \frac{\hat{y}_2(k) - y(k)}{\hat{y}_2(k) - \hat{y}_1(k)}$$

Plugging into Eq. (5.17) and noting that $\alpha_2(k) = 1 - \alpha_1(k)$, one can further derive

$$\begin{aligned}
\hat{y}(k) &= \alpha_1(k)\hat{y}_1(k) + \alpha_2(k)\hat{y}_2(k) \\
&= \alpha_1(k)\hat{y}_1(k) + (1-\alpha_1(k))\hat{y}_2(k) \\
&= \hat{y}_2(k) + (\hat{y}_1(k) - \hat{y}_2(k))\hat{\alpha}_1(k) \\
&= \hat{y}_2(k) + (\hat{y}_1(k) - \hat{y}_2(k))\frac{\hat{y}_2(k) - y(k)}{\hat{y}_2(k) - \hat{y}_1(k)} \\
&= y(k)
\end{aligned}$$

Thus, if α converges, the estimate of the multiple models $\hat{y}(k)$ converges to the true value $y(k)$.

Algorithm 2 (Multiple Model adaptation)

1. *Set the initial coefficients* $\alpha_1 = \alpha_2 = 0.5$ *at time* t_0. *Choose* β *as* $0 < \beta < 1$.
2. *At time instant* k, *compute the output of both the two local models* $\hat{y}_1(k), \hat{y}_2(k)$; *compute the MM output* $\hat{y}(k)$ *by Eq.* (5.17); *and compute* $e(k), e_2(k), \tilde{E}(k)$.
3. *Compute the updating rule of* (5.15) *and obtain* α_1.
4. *Obtain* $\alpha_2(k)$ *by* $\alpha_2(k) = 1 - \alpha_1(k)$.

In practice, to ensure the condition of (5.18) holds, one can first simulate the two local models or conduct experiments, and calculate the maximum difference between their outputs, denoted as $\tilde{E}_{max}$. Then, β can be chosen as $\beta < \frac{2}{\tilde{E}^2_{max}}$.

5.5. MODELING AND SIMULATION OF THE EGFR SIGNALING NETWORK BY NARX AND MULTIPLE-MODEL PREDICTION

The NARX-PSO method and the multiple model approach proposed in the previous sections will be applied to model and simulate the EGFR signaling network in this section. In 5.1, PSO and RWBS are compared in modeling the subnets in the EGFR signaling network. In 5.2, the performance of MM is studied. Moreover in 5.5.3, a cascade model

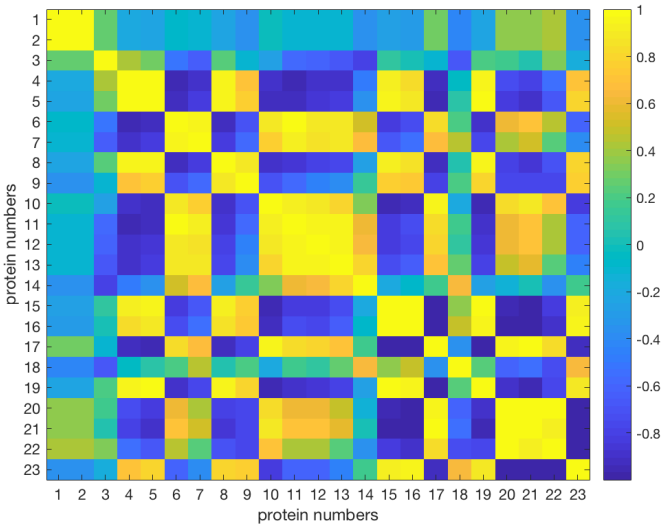


Figure 5.2: Correlation map of the time-series data from the concentrations of the proteins of the EGFR signaling network

connecting 2 subnets is proposed and tested. Modeling the total phosphorylated proteins is finally investigated and tested in 5.5.4.

The simulations in this section are performed in the MATLAB R2017b environment on a computer with an Intel Core i5 2.70 GHz CPU and 8.00GB RAM.

5.5.1. SUBNET MODELING

The EGF network as defined in Fig. 5.1 contains 25 different reactions, and is hence a rather complicated structure. For simplicity, we number all the proteins from 1 to 23 as listed in Table A.1 in the Appendix. A.3. To ease the modeling, we first apply a "divide-and-conquer" approach, and divide the network into several subnets. The dividing criteria is the cross-correlation among the 23 different proteins. The cross correlation is computed as follows. First, we simulated the ODEs of the concentrations of the EGFR signaling network by the "ODE45" function in MATLAB. Then, a time-series data with 600 samples are generated for each of the 23 proteins. The time interval was set to 0.2 minute. After simulating for 120 minutes, the time series hence contain 600 samples.The cross correlation is plotted in Fig. 5.2. On the other hand, different initial concentrations of 2, 3, 4, 5, 6, 10, 15, 20nM of EGF are set in the simulations. These eight simulations give rise to eight different datasets, respecting labeled as "EGF=2nM", $\cdots$, "EGF=20nM".

According to the cross-correlation in Fig. 5.2, all the 23 proteins are correlated with each other. The correlation between RP (protein 5) and Shp (protein 15) is more than 0.9, which is obviously highly correlated. Hence, we can build a subnet between RP and Shp. In such way, we can build subnet models between any two proteins in this network that are highly correlated. For instance, another example to be studied is the subnet from

Table 5.1: Performance of the NARX models tested by seven datasets with kernels selected respectively by RWBS and PSO.

testing data	subnet 1		subnet 2	
	RWBS	PSO	RWBS	PSO
EGF=2	0.2019	0.1980	1.3490	1.3765
EGF=3	0.1488	0.0987	0.58638	0.6038
EGF=4	0.0665	0.0503	0.3169	0.3178
EGF=5	0.0379	0.0336	0.1997	0.2010
EGF=6	0.0301	0.0270	0.1395	0.1391
EGF=10	0.0253	0.0219	0.0561	0.0549
EGF=15	0.0249	0.0213	0.0299	0.0299

Shp to Sh-G-S.

In modeling the subnets from RP to Shp and Shp to Sh-G-S, we compare the kernel selection methods respectively by the PSO and RWBS algorithm. The parameters of RWBS algorithm [22] are chosen the same as those used in [23] i.e.$Q = 50$, $T = 600$. The PSO parameters are $Q = 50$ and $T = 5$, which can yield comparable model accuracy as the RWBS algorithm, as will be shown later in the simulation. The weighting factors in Eq. (5.3) are set to $w = 0.6, c1 = 0.6, c2 = 0.5$ empirically. Because the procedures in the outer loop of both the algorithms are almost the same, we only compare the time cost of running the PSO and RWBS inner loop. The training process uses dataset "EGF=20nM", while the testing process uses the other seven.

The first subnet to be modeled is from RP to Shp, we take the concentration of RP as input denoted as $u_{RP}(k)$, and the concentration of Shp as output denoted as $y_{Shp}(k)$. Take $n_u = 3, n_y = 1$. The NARX model can therefore be written as

$$y_{Shp}(k) = f(u_{RP}(k-1), u_{RP}(k-2), u_{RP}(k-3), y_{Shp}(k-1))$$

The training datasets are shown in Fig. 5.3.

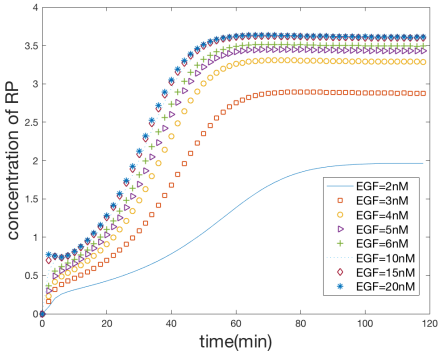
For this problem, the average time cost of running the RWBS algorithm was about 4.3 seconds; while the average time cost of the PSO algorithm was about 1.3 seconds. Both of them ended up with 5 to 8 best Gaussian RBF kernels. The evolution of the cost function in both cases is shown in Fig. 5.4.

Another subnet model takes the concentration of Shp as input denoted as $u_{Shp}(k)$, and the concentration of Sh-G-S as output denoted as $y_{Sh-G-S}(k)$. Also take $n_u = 3, n_y = 1$. The NARX model can be written as

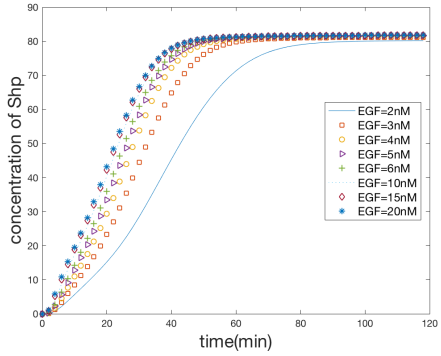
$$y_{Sh-G-S}(k) = f(u_{Shp}(k-1), u_{Shp}(k-2), u_{Shp}(k-3), y_{Sh-G-S}(k-1))$$

The average time cost of running RWBS algorithm was about 4.5seconds; while he average time cost of PSO algorithm was about 1.4 seconds. Both of them can result in a 6 to 8 best Gaussian RBF kernels. The model accuracy of the two subnets are compared in Table 5.1. The accuracy by both kernel selection methods are similar.

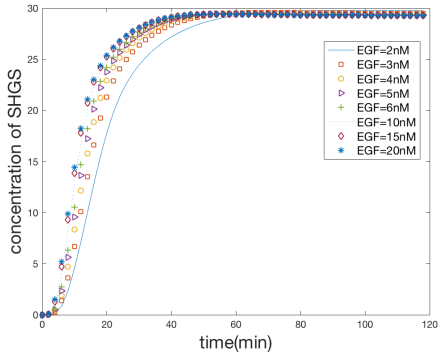
Whereas, the time cost of the PSO algorithm is much lower than that of the RWBS.This can be attribute to the fact that the RWBS needs to repeat for multiple times with random initializations to convert a local optimizer to a global one [22]. Therefore, to find



(a)

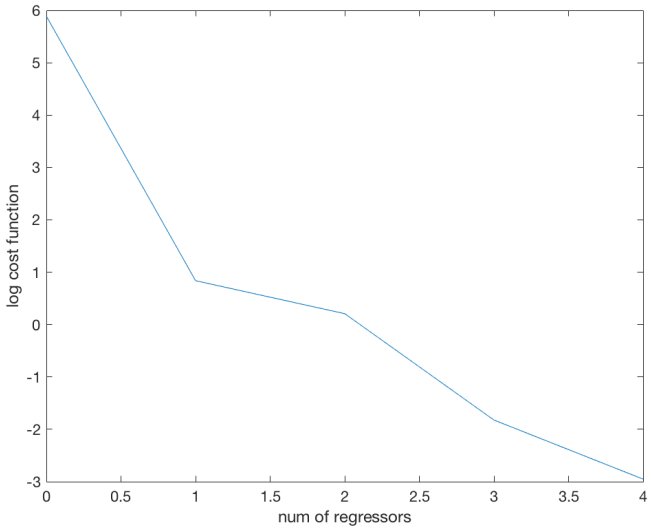


(b)

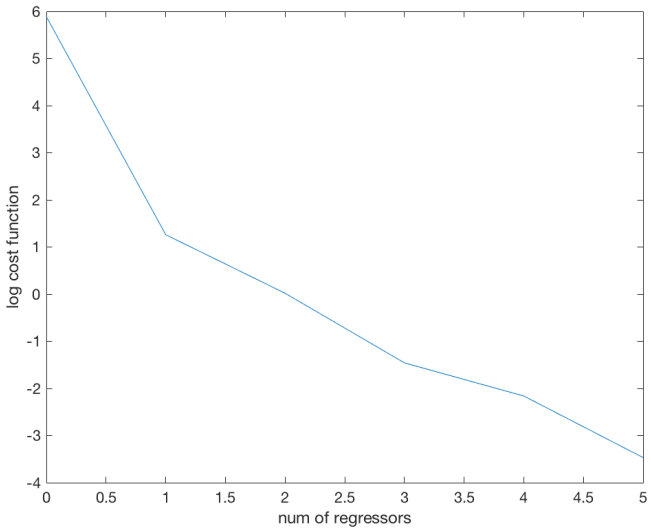


(c)

Figure 5.3: The time series of the concentration of the proteins simulated at eight different initial concentrations of EGF: (a) RP; (b) Shp; (c) Sh-G-S.



(a)



(b)

Figure 5.4: The evolution of the cost function during the kernel selection: (a) selection by RWBS; (b) selection by PSO.

Table 5.2: The average MSE of resulted from the local models and the multiple model approach, where the average MSE is the average MSE errors from testing with the six datasets generated from "EGF=3nM" to "EGF=15nM"

average MSE	local model 1	local model 2	multiple model
subnet 1	0.7688	0.2265	0.0245
subnet 2	0.6586	0.2110	0.0263

each kernel center, the RWBS algorithm needs to repeats for several times, which significantly increases the computational load.

On the other hand, while testing with the data generated at low EGF concentrations, the MSE errors were relatively larger than using the other datasets, because of the larger differences in the initial EGF concentrations therein than that in the training data.

5.5.2. MULTIPLE MODEL ADAPTATION

According to [2], the concentrations of the phosphorylated proteins are higher along the entire time sequence at higher EGF concentration than those at lower ones. We hence need two local models respectively corresponding to a high EGF concentration and a low each individual. The two local models were trained with the datasets corresponding to these two different initial EGF concentrations. For example, in subnet 1, the model is constructed between the concentration of RP and the concentration of Shp. The output of local model 1 is modeled as

$$y^1_{Shp}(k) = f(u^1_{RP}(k-1, u^1_{RP}(k-2), u^1_{RP}(k-3), y^1_{Shp}(k-1))$$

The output of local model 2 is modeled as

$$y^2_{Shp}(k) = f(u^2_{RP}(k-1), u^2_{RP}(k-2), u^2_{RP}(k-3), y^2_{Shp}(k-1))$$

Finally, the outputs from two local models are convexly combined as described in Eq. (5.10), which leads to the predicted outputs.

From the results listed in Table 5.2, neither of the local models can fit the testing data well. Furthermore, the more difference between the initial concentration of EGF of the training set and the testing set, the higher MSE error resulted from each individual local model. The results from the multiple model prediction are clearly much more accurate.

5.5.3. CASCADE MODEL

Due to the large number of proteins and the complexity of the network structure, the relationship between two proteins that are "far from" each other, in terms of their connections via multiple edges may be too weak to lead to an accurate model. Therefore, in this work, we propose to build cascade models, which connect more than one subnet models in series, to link the proteins along a relatively long chain with stronger mathematical relations. The cascade model takes the output of the former subnet as input, and in turn feeds its output to the next subnet. For example, both the output of subnet 1 and the input of subnet 2 is the concentration of ShP. So we can connect these two subnets together. The structure is described in Fig. 5.5.

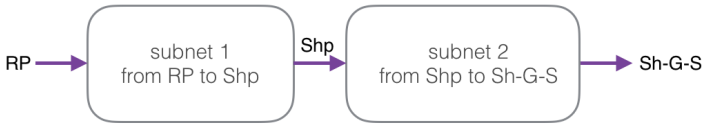


Figure 5.5: The structure of the cascade model.

In this cascade model, the input is the concentration of RP, the output is the concentration of Sh-G-S. The average MSE error resulted from this cascade model is 0.0226. Whereas, directly modeling from the concentration of RP to that of Sh-G-S, i.e.

$$y_{Sh-G-S}(k) = f(u_{RP}(k-1), u_{RP}(k-2), u_{RP}(k-3), y_{Sh-G-S}(k-1))$$

gave rise to an average MSE of 0.0751, which is three times larger than the cascade model. As an example, the predictions by the cascade model, when the initial EGF concentration was 4nM, is shown in Fig. 5.6.

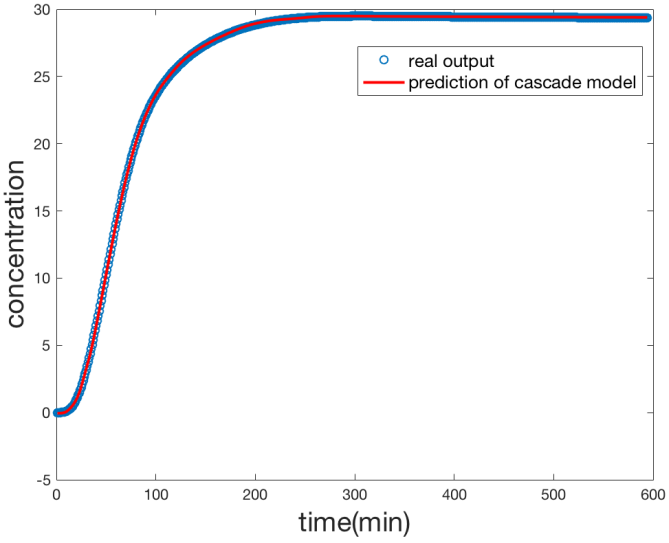


Figure 5.6: Prediction by the cascade model, when the initial EGF concentration was 4nM.

5.5.4. MODEL OF THE TOTAL PHOSPHORYLATED PROTEINS

In cell experiments, the concentrations of the tyrosine phosphorylation of specific target proteins, such as Shc, Grb and EGFR, can be measured. This measurement is easier than measuring individual proteins. The total concentrations of the target proteins are defined in [2]. Here, we take the total concentrations of phosphorylated Shc as an example.

Table 5.3: Performance of the MM-based NARX in modeling the total concentration of Shc.

	local model 1	local model 2	multiple model
EGF=3	3.5968	1.05170	0.0569
EGF=4	6.4781	0.2770	0.0429
EGF=5	7.7445	0.1667	0.0470
EGF=6	8.3547	0.1384	0.0496
EGF=10	9.1945	0.1262	0.0528
EGF=15	9.4905	0.1256 6	0.0526

The concentration of total phosphorylated Shc is defined as

$$y_{t,Shc} = C_{R-ShP} + C_{R-Sh-G} + C_{R-Sh-G-S} + C_{ShP} + C_{Sh-G} + C_{Sh-G-S},$$

where the subscript "t" means the total concentration of phosphorylated Shc. The time-series of $y_{t,Shc}$ is shown in Fig. 5.7. The NARX model takes the concentration of the total phosphorylated proteins as the output denoted as $y_{t,Shc}(k)$, and the concentration of R as the input denoted as $u_R(k)$, which takes the following form

$$y_{t,Shc}(k) = f(u_R(k-1), u_R(k-2), u_R(k-3), y_{t,Shc}(k-1)).$$

The testing results from the six different datasets, respectively corresponding to six different initial EGF concentrations, are listed in Table 5.3. As an example, the predicted output of $y_{t,Shc}$ when the initial EGF concentration was 4nM is plotted in Fig. 5.8. It can be seen that although the output trajectories become more chaotic, the multiple-model approach still gave a more accurate result than the two local models did.

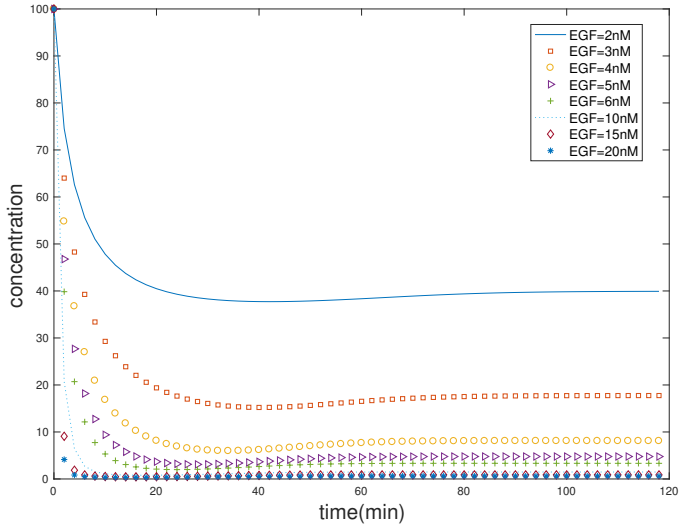
5.6. CONCLUSION

In this paper, we have proposed a modeling scheme with kernel-based NARX structure, together with a PSO-based kernel selection method. This scheme has been applied to model the EGFR signaling network in mammalian cells to handle the nonlinear dynamics therein. Moreover, a multiple model adaptation method has also been developed to handle the chaotic nature of this complicated biochemical process. Three simulation case studies have been conducted; i.e. on small scale subnet models with proteins that are highly correlated, on larger scale cascade models with proteins exhibiting weak correlation, and on the total amount of phosphorylated proteins. All the simulation results have verified the model accuracy.

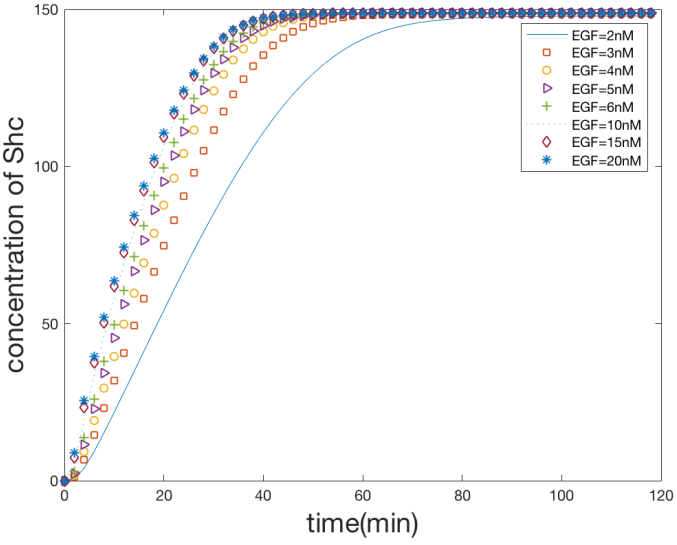
As one future direction, the models will be identified from real-life experimental data by conducting cell experiments. Another direction is to model the effect of exogenous stimuli. e.g. the effect of medicine or physical factors like photons, to the EGFR signaling network in mammalian cells.

REFERENCES

[1] Y. Yarden and J. Schlessinger, *Epidermal growth factor induces rapid, reversible aggregation of the purified epidermal growth factor receptor,* Biochemistry **26**, 1443 (1987).



(a)



(b)

Figure 5.7: The time series of the concentrations of the proteins simulated at eight different initial concentrations of EGF. (a) R (b) total phosphorylated Shc.

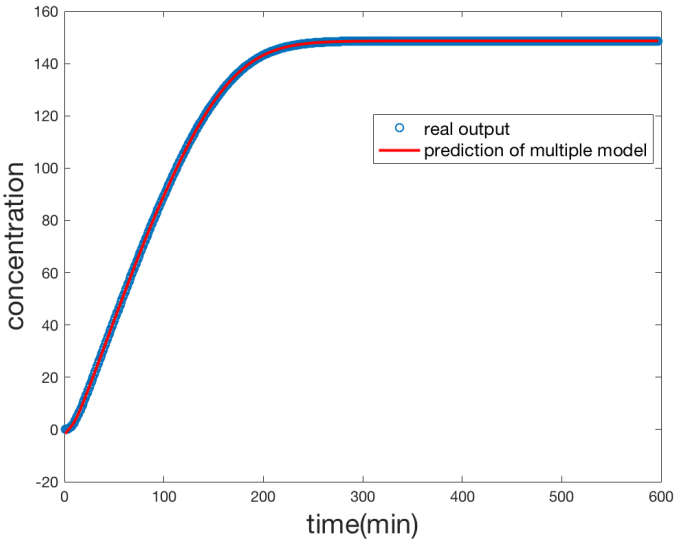


Figure 5.8: Prediction of the total phosphorylated Shc, when the initial EGF concentration was 4nM.

[2] B. N. Kholodenko, O. V. Demin, G. Moehren, and J. B. Hoek, *Quantification of short term signaling by the epidermal growth factor receptor,* Journal of Biological Chemistry **274**, 30169 (1999).

[3] G. Moehren, N. Markevich, O. Demin, A. Kiyatkin, I. Goryanin, J. B. Hoek, and B. N. Kholodenko, *Temperature dependence of the epidermal growth factor receptor signaling network can be accounted for by a kinetic model,* Biochemistry **41**, 306 (2002).

[4] D. Zhou and P. Frank, *A real-time estimation approach to time-varying time delay and parameters of NARX processes,* Computers and Chemical Engineering **23**, 1763– (2000).

[5] D. Zhai, D. Rollins, N. Bhandari, and H. Wu, *Continuous-time hammerstein and wiener modeling under second-order static nonlinearity for periodic process signals,* Computers and Chemical Engineering **31**, 1 (2006).

[6] J. Gan, S. Parulekar, and A. Cinar, *Development of a recursive time series model for fed-batch mammalian cell culture,* Computers and Chemical Engineering **109**, 289 (2018).

[7] W. Lee, N. Na, K. Kim, C. Lee, Y. Lee, and J. Lee, *NARX modeling for real-time optimization of air and gas compression systems in chemical processes,* Computers and Chemical Engineering **115**, 262– (2018).

[8] J. Dong, *Recursive NARX model identification of nonlinear chemical processes with matrix invertibility analysis,* Computers and Chemical Engineering **121**, 574 (2019).

[9] S. Chen, C. Cowan, and P. Grant, *Orthogonal least squares learning algorithm for radial basis function networks,* IEEE Transactions on Neural Networks **2**, 302 (1991).

[10] B. Scholkopf and A. Smola, *Learning with Kernels* (MIT Press, 2002).

[11] J. Suykens, C. Alzate, and K. Pelckmans, *Primal and dual model representations in kernel-based learning,* Statistics Surveys **4**, 148 (2010).

[12] R. Tóth, P. Heuberger, and P. V. den Hof, *Asymptotically optimal orthonormal basis functions for LPV system identification,* Automatica **45**, 1359 (2009).

[13] S. Chen, X. Hong, C. J. Harris, and P. M. Sharkey, *Sparse modeling using orthogonal forward regression with press statistic and regularization,* IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) **34**, 898 (2004).

[14] X. Hong, P. Sharkey, and K. Warwick, *Automatic nonlinear predictive model-construction algorithm using forward regression and the press statistic,* IEE Proceedings-Control Theory and Applications **150**, 245 (2003).

[15] S. Chen, Y. Wu, and B. Luk, *Combined genetic algorithm optimization and regularized orthogonal least squares learning for radial basis function networks,* IEEE Transactions on Neural Networks **10**, 1239 (1999).

[16] J. Kennedy and R. Eberhart, *Particle swarm optimization,* Proceedings of IEEE International Conference on Neural Networks (1995).

[17] Y. Shi and R. Eberhart, *A modified particle swarm optimizer,* Proceedings of IEEE International Conference on Evolutionary Computation (1998).

[18] M.-H. Golbon-Haghighi, H. Saeidi-Manesh, G. Zhang, and Y. Zhang, *Pattern synthesis for the cylindrical polarimetric phased array radar (CPPAR),* Progress In Electromagnetics Research **66**, 87 (2018).

[19] A. Ferrari, S. Gutierrez, and E. C. Biscaia Jr, *Development of an optimal operation strategy in a sequential batch reactor (SBR) through mixed-integer particle swarm dynamic optimization (PSO),* Computers & Chemical Engineering **34**, 1994 (2010).

[20] C. O. Ourique, E. C. Biscaia Jr, and J. C. Pinto, *The use of particle swarm optimization for dynamical analysis in chemical processes,* Computers & Chemical Engineering **26**, 1783 (2002).

[21] C. Ou and W. Lin, *Comparison between pso and ga for parameters optimization of pid controller,* in *2006 International conference on mechatronics and automation* (IEEE, 2006) pp. 2471–2475.

[22] S. Chen, X. Wang, and C. J. Harris, *Experiments with repeating weighted boosting search for optimization signal processing applications,* IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) **35**, 682 (2005).

[23] S. Chen, X. Wang, and C. J. Harris, *Narx-based nonlinear system identification using orthogonal least squares basis hunting,* IEEE Transactions on Control Systems Technology **16**, 78 (2007).

[24] N. Indera, I. Yassin, A. Zabidi, and Z. Rizman, *Non-linear autoregressive with exogeneous input (NARX) bitcoin price prediction model using pso-optimized parameters and moving average technical indicators,* Journal of Fundamental and Applied Sciences **9**, 791 (2017).

[25] M. Mohamad, I. Yassin, A. Zabidi, M. N. Taib, and R. Adnan, *Comparison between pso and ols for NARX parameter estimation of a dc motor,* in *2013 IEEE Symposium on Industrial Electronics & Applications* (IEEE, 2013) pp. 27–32.

[26] Y. Wang, *A new concept using lstm neural networks for dynamic system identification,* American Control Conference , 5324 (2017).

[27] K. S. Narendra and Y. Wang, *Simulation studies of feed-forward learning schemes with multiple models,* Yale technical report , 1603 (2016).

[28] R. Murray-Smith and T. Johansen, *Multiple model approaches to nonlinear modelling and control* (CRC press, 1997).

[29] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, *Interacting multiple model methods in target tracking: a survey,* IEEE Transactions on aerospace and electronic systems **34**, 103 (1998).

[30] H. Wang, P. Li, F. Gao, Z. Song, and S. X. Ding, *Kernel classifier with adaptive structure and fixed memory for process diagnosis,* AIChE journal **52**, 3515 (2006).

6

ANALYSIS AND DATA-BASED MODELING OF THE PHOTOCHEMICAL REACTION DYNAMICS OF THE INDUCED SINGLET OXYGEN IN LIGHT THERAPIES

Objective: *The macroscopic singlet oxygen (MSO) model for quantifying the light-induced singlet oxygen ($^1O_2$) always contain a set of nonlinear dynamic equations and therefore are generally difficult to be applied. This work was devoted to analyze and simplify this dynamic model.* Methods: *Firstly, the nonlinearity of the MSO model was analyzed with control theory. The conditions, under which it can be simplified to a linear one, were derived. Secondly, in the case of ample triplet oxygen concentration, a closed-form exact solution of the $^1O_2$ model was further derived, in a nonlinear algebraic form with only four parameters that can be easily fitted to experimental data. Finally,* in vitro *experiments of anti-fungal light therapies were conducted, where the fungi,* Candida albicans, *were irradiated respectively by the 385, 405, 415, and 450nm wavelength light. The singlet oxygen concentration levels in the fungi were measured, and then used to fit the developed models.* Results: *The parameters of the closed-form exact solution were estimated from both the simulated and the measured experimental data. Based on this model, a functional relationship between the photon energy, fluence rate and singlet oxygen concentration was also established. The fitting accuracy of this model to the data was satisfactory, which therefore demonstrates the effectiveness of the proposed modeling techniques.* Conclusion: *The results from simulating the closed-form model indicate that the photon energy*

within the range of either 2.7 ~ 2.8 eV or 3.0 ~ 3. 2eV (388 ~ 413nm or 443 ~ 459nm in wavelength) is more effective in generating singlet oxygen in the fungi studied in this work. Significance: *It is the first attempt of applying control theory to analyze the photochemical reaction dynamics of light therapies in terms of their nonlinearity. The proposed modeling techniques also offer opportunities for determining the light dosages in treating fungal infection diseases, especially those on the surface tissues of human body.*

6.1. INTRODUCTION

Photodynamic therapies (PDT) have been investigated as non-antibiotic alternatives for treating localized infectious diseases due to their rapid action and lack of drug resistance [1, 2]. Similar to PDT, blue light therapies that rely only on the endogenous pigments, i.e., porphyrins and flavins, of the pathogens are also effective and even safer to use. The inactivation effect of blue light has been proven on the Gram-positive/negative bacteria [3], mycobacterium [4], and algae [5]. Likewise, anti-fungal blue light (ABL) has been widely studied as a new treatment approach of fungal infections. For instance, the effect of 415nm LED light on *C. albicans* was studied and verified both *in vitro* and *in vivo* [6]. Blue light at 405nm was also previously shown to be effective in reducing the seeding of exposed biofilms [7].

As a widely accepted hypothesis, the mechanisms of the anti-fungal effect of PDT and ABL are that the photons from light can excite either the exogenous photosensitizer (PS) in the former case or the endogenous pigments in the latter, which in turn produces reactive oxygen species (ROS) that affect cells [8, 9]. ROS are highly reactive chemical molecules formed due to the electron acceptability of O_2. Examples of ROS include peroxides, superoxide, hydroxyl radicals, singlet oxygen [10], and alpha-oxygen [11]. According to a previous study [12], singlet oxygen usually accounts for 80% of all the ROS induced by light; while hydroxyl radicals and other types of ROS take the remaining 20%. Furthermore, singlet oxygen is a precursor of most other ROS, and it also becomes involved in the propagation of oxidative chain reactions [13].

Thus, there is greater interest in the quantum yield of singlet oxygen, whose cumulative dose has been used to determine the tissue damage caused by PDT [14–16]. Based on the mechanism of PDT, some first-principle models have been proposed to explore the dynamic changes in singlet oxygen concentration [17–19], which are highly nonlinear models. A set of coupled differential equations was used to describe this nonlinear process, including seven Michaelis-Menten-type equations corresponding to seven different state variables [19], i.e., the concentration of PS and various types of oxygen and the ROS acceptors excluding the PS molecules. Moreover, Monte Carlo simulations were also combined with these types of kinetic models, and were proven to be effective in simulating the light transport in biological tissues [20–22].

These first-principle models can precisely fit the dynamic changes in the singlet oxygen concentrations. However, the parameters are related to the PS characteristics, and they should be determined from dedicated experiments. For instance, in a previous study [23], up to 21 parameters were required to describe the process related to the PS. A classical macroscopic time-scale singlet oxygen (MSO) model [18, 19] was simplified into three differential equations [21, 22] with some reasonable assumptions according to the reality. However, 6 parameters and 3 states remained. This simplified model was

verified in a previous study [24], in which the concentrations of PS and triplet oxygen concentration were measured. The parameters were taken as the empirical values because of the difficulty in re-estimating them. However, although ABL are believed to be caused by the PSs that naturally exist in fungal cells, whose types and amounts are usually unknown, modeling by first-principles becomes even more challenging than modeling PDT. To accommodate these modeling challenges, data-driven modeling methods based on time-series have been investigated [25–27]. However, there is still no attempt made to ease the model form and the parameter estimation in the classical MSO models, in a similar data-driven fashion based on measured time-series data.

In this study, we first examined the nonlinear dynamics of the MSO model, and then analyzed the conditions under which it can be approximated by a linear model. Based on the standard system linearization method, a linearized model was developed and applied to simulate the concentration of singlet oxygen. We subsequently solved the differential equations in the MSO model for the case of an abundant oxygen supply, and obtained an analytical solution of the concentration of singlet oxygen. In experiments, the singlet oxygen levels are measured from *C. albicans* under the irradiation of four LED light sources with wavelengths of 385, 405, 415, and 450nm up to 30 minutes. Here, we chose *C. albicans* to study because it is one of the most common fungal pathogens [28], and it always infects membranes such as those of the mouth, vagina, and intestines in humans [29, 30]. The fitting accuracy of the analytical solution of the MSO model to the measured singlet oxygen concentrations also turned out to be satisfactory.

6.2. METHODS

6.2.1. MODELING METHODS

PRELIMINARIES OF THE FIRST-PRINCIPLE MODEL

A classical macroscopic time-scale singlet oxygen (MSO) model is based on the mechanism of PDT with the included rate equations in Michaelis-Menten form [18, 19]. With some reasonable assumptions according to the reality, this model can be simplified into three differential equations [21, 22].

$$\frac{d[S_0]}{dt} = -\xi\sigma\frac{\phi([S_0]+\delta)[^3O_2]}{[^3O_2]+\beta}[S_0] \tag{6.1a}$$

$$\frac{d[^3O_2]}{dt} = -\xi\frac{\phi[S_0]\,[^3O_2]}{[^3O_2]+\beta} + g\cdot(1-\frac{[^3O_2]}{[^3O_2]_0}) \tag{6.1b}$$

$$\frac{d[^1O_2]}{dt} = f\cdot\xi\frac{\phi[S_0][^3O_2]}{[^3O_2]+\beta} \tag{6.1c}$$

where $[S_0]$ denotes the concentration of PS; $[^3O_2]$, $[^1O_2]$ denote the concentration of triplet and singlet oxygen, respectively; and $[^3O_2]_0$ denotes the initial concentration of triplet oxygen. The included parameters are light fluence rate ϕ [mW], specific oxygen consumption rate ξ [cm^2/mW/s], the specific photobleaching ratio σ [μM^{-1}], the oxygen quenching threshold β [μM], the low concentration correction δ [μM], the fraction of $^1O_2$ interacting with target f, and the oxygen supply rate g [μM/s]. In an ample oxygen supply, the fraction f can be set as 1. Let g to be the external input, the nonlinearity of

$$\boldsymbol{A} = \begin{bmatrix} -\xi\sigma\phi\frac{[^3O_2]}{[^3O_2]+\beta}(2[S_0]+\delta) & -\xi\sigma\phi([S_0]+\delta)[S_0](\frac{1}{\beta+[^3O_2]} - \frac{[^3O_2]}{(\beta+[^3O_2])^2}) & 0 \\ -\xi\phi\frac{[^3O_2]}{[^3O_2]+\beta} & -g/[^3O_2]_0 - \xi\phi[S_0](\frac{1}{\beta+[^3O_2]} - \frac{[^3O_2]}{(\beta+[^3O_2])^2}) & 0 \\ f\xi\phi\frac{[^3O_2]}{[^3O_2]+\beta} & f\xi\phi[S_0](\frac{1}{\beta+[^3O_2]} - \frac{[^3O_2]}{(\beta+[^3O_2])^2}) & 0 \end{bmatrix} \tag{6.4}$$

model (6.1) can be examined by the poles and zeros of its linearized model at different equilibrium points. If $\boldsymbol{x} = \begin{bmatrix}[S_0] & [^3O_2] & [^1O_2]\end{bmatrix}^T$ is denoted as the state vector, and $u = g$ as the control input of the nonlinear system, then Eq. (6.1) can be rewritten in a state space form as

$$\frac{d\boldsymbol{x}}{dt} = h(\boldsymbol{x}, u), \tag{6.2}$$

where $h(\cdot)$ is a nonlinear vector-valued function.

ANALYSIS OF THE FIRST-PRINCIPLE MODEL OF PDT

To analyze Eq. (6.1), first note that if a nonlinear system operates around an equilibrium point, i.e. around a configuration where the system is at rest, then it is possible to study the behavior of the system in the neighborhood of such a point, where its dynamics can be reasonably approximated as a linear model. Based on the standard linearization method of a nonlinear system [31], we can linearize Eq. (6.2) as

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}u, \tag{6.3}$$

where $\boldsymbol{A}$ is defined as $\boldsymbol{A} = \frac{\partial h}{\partial \boldsymbol{x}}\Big|_{\boldsymbol{x}^*,u^*}$; $\boldsymbol{B}$ is defined as $\boldsymbol{B} = \frac{\partial h}{\partial u}\Big|_{x^*,u^*}$; and $\boldsymbol{x}^*, u^*$ denotes an equilibrium point. Thus, the linearized matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ can be expressed as Eq. (6.4) on the top of next page and Eq. (6.5).

$$\boldsymbol{B} = \begin{bmatrix}0 & 1 - [^3O_2]/[^3O_2]_0 & 0\end{bmatrix}^T \tag{6.5}$$

Because the main objective is to model and track the concentration of $^1O_2$, an output equation can be added to Eq. (6.3), which leading to a standard state-space form:

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}u \tag{6.6a}$$

$$\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x}. \tag{6.6b}$$

Here, as a reminder, the state vector is $\boldsymbol{x} = \begin{bmatrix}[S_0] & [^3O_2] & [^1O_2]\end{bmatrix}^T$. The output matrix $\boldsymbol{C} = \begin{bmatrix}0 & 0 & 1\end{bmatrix}$, since the third dimension of the state vector is $[^1O_2]$.

To examine the level of the nonlinearity of this dynamic process, one can study the variation of the poles at different equilibrium points. It is well known in the control theory literature that the poles and zeros of a system determine its dynamic behavior in response to exogenous stimuli [31]. First, note that for the state-space model (6.3), the poles are the eigenvalues of system matrix $\boldsymbol{A}$ [31], which are denoted as λ. Because $\boldsymbol{A} \in \mathbb{R}^{3\times3}$, it has three eigenvalues. Based on the standard linear algebra, one eigenvalue

is 0, because the third column of $\boldsymbol{A}$ are all 0's; while the other two can be obtained by solving the following characteristic equation:

$$(\lambda - a_{11})(\lambda - a_{22}) - a_{12} \cdot a_{21} = 0, \tag{6.7}$$

where a_{ij} denotes the element on the i-th row and j-th column of $\boldsymbol{A}$. Therefore, three eigenvalues can be written as:

$$\begin{aligned}
\lambda_1 &= 0 \\
\lambda_2 &= \tfrac{1}{2}\left[a_{11} + a_{22} - \sqrt{(a_{11}+a_{22})^2 - 4(a_{11}\cdot a_{22} - a_{12}\cdot a_{21})}\right] \\
\lambda_3 &= \tfrac{1}{2}\left[a_{11} + a_{22} + \sqrt{(a_{11}+a_{22})^2 - 4(a_{11}\cdot a_{22} - a_{12}\cdot a_{21})}\right].
\end{aligned}$$

Under the real conditions of photofrin-mediated PDT [21, 22], the term σ is far less than the other parameters, i.e., $\sigma = 9{\cdot}10^{-5}\ \mu\text{M}^{-1}$, $\xi = 3.7{\cdot}10^{-3}\ \text{cm}^2/\text{mW/s}$, $\delta = 33\ \mu\text{M}$, $\beta = 11.9\ \mu\text{M}$. This leads to $a_{11}, a_{12} \approx 0$, which contains σ. Thus, the term $a_{11}\cdot a_{22} - a_{12}\cdot a_{21} \approx 0$. After substituting this into the equation above, the eigenvalues can be simplified to

$$\begin{cases} \lambda_1 & = 0 \\ \lambda_2 & \approx 0 \\ \lambda_3 & \approx a_{11} + a_{22} \end{cases}. \tag{6.8}$$

Obviously, there is a pole at the origin (λ_1) and two negative real poles. However, the analytic solutions of the zeroes are not as straightforward as those of the poles. Whereas, according to the simulation results to be presented later, the linearized models only contain one zero. Therefore, the linearized model can also be written in a transfer function form, i.e.,

$$H(s) = K \cdot \frac{s - z}{s(s-\lambda_2)(s-\lambda_3)} \tag{6.9}$$

$$\approx \frac{K}{(s-\lambda_2)(s-\lambda_3)}, \tag{6.10}$$

where K and z are respectively the gain and zero. According to the simulations to be presented later, the order of magnitude of z is only -5 in all the linearized systems. This trivial zero always cancels the pole at the origin, which approximately leads to a second order transfer function (6.10) by standard minimal realization theory. This will also be illustrated in the root-locus plots of the linearized systems at different equilibrium points in Sec. 6.3.1. Consequently, the transfer function (6.10) contains a dominant pole (λ_2) and a non-dominant one (λ_3).

Based on the locations of the zero and poles of the linearized systems, we mainly need to analyze λ_2 and λ_3. Besides, due to the simpler form of λ_3, we will now focus on analyzing λ_3 to investigate the nonlinearity of the MSO process. First, substituting the corresponding terms of Eq. (6.4) into λ_3 as follows, we obtain:

$$\lambda_3 = -\xi\sigma\phi\frac{[^3O_2](2[S_0]+\delta)}{[^3O_2]+\beta} - \frac{g}{[^3O_2]_0} - \frac{\xi\phi\beta[S_0]}{(\beta+[^3O_2])^2} \tag{6.11}$$

Here, λ_3 is a negative real pole, which defines an exponentially decaying component in the dynamic response, i.e., $e^{\lambda_3 t}$.

In this MSO model, the initial concentration of triplet oxygen and the oxygen supply rate can be artificially adjusted by the oxygen content of the environment where the cells are exposed. If the external environment is oxygen-rich, one can expect that the concentration of triplet oxygen, i.e., $[^3O_2]$, is far greater than β, and thus, the limit of λ_3 is simplified to:

$$\lim_{[^3O_2]\to+\infty} \lambda_3 = -\xi\sigma\phi(2[S_0]+\delta) \tag{6.12}$$

Because ξ, σ, and ϕ are constant parameters, the concentration of PS, i.e., $[S_0]$, becomes the only factor that can change the third pole, which in turn affects the dynamic response of the system. As will be detailed later in the simulation results, there is actually little change in $[S_0]$ during the whole process. Thus, the limit (6.12) is approximately a constant. Similarly, the same conclusion can be drawn on $\lim_{[^3O_2]\to+\infty}\lambda_2$, which are also supported by the simulation results in Sec. 6.3.1.

Taking into account that λ_2 and λ_3 are approximately constants, the MSO model can thus be approximated by a linear time invariant model, in the case of abundant oxygen supply. Thus, one can use the linearized model (6.6) to simulate the concentration of light-induced singlet oxygen $[^1O_2]$.

It is worth mentioning that the linearized model (6.6) is a second-order linear-time-invariant (LTI) model after the pole-zero cancelation, containing two poles. In comparison, another LTI model [26] is a first-order model with only one pole, whose value is close to the dominant pole λ_2.

ANALYTICAL SOLUTION OF $[^1O_2]$ WITH ABUNDANT $^3O_2$

Although the approximated linear model (6.6) provides additional insights into the dynamic behavior of the MSO process, it can also lead to large prediction errors in the long-term trajectories of $[^1O_2]$ due to the approximations made for both the zeros and poles. This further motivates the derivation of an exact solution of the differential equations in (6.1), in the case of abundant $^3O_2$.

Based on the situation of the abundant oxygen supply, $[^3O_2]$ is much greater than β. With this condition, $\frac{[^3O_2]}{[^3O_2]+\beta}$ can approximate to a constant as

$$\lim_{[^3O_2]\to+\infty} \frac{[^3O_2]}{[^3O_2]+\beta} = 1. \tag{6.13}$$

It is difficult to derive a closed-form solution of $[^1O_2]$ from the linearized model of Eq. (6.3), which contains three states. Neglecting the variation in the concentration of $^3O_2$ and the approximation in Eq. (6.13), we can eliminate Eq. (6.1b) and derive the following second-order dynamics:

$$\frac{d[S_0]}{dt} = a([S_0]+\delta)[S_0] \tag{6.14a}$$

$$\frac{d[^1O_2]}{dt} = c[S_0] \tag{6.14b}$$

where $a = -\xi\sigma\phi$ and $c = f\cdot\xi\phi$. Note that, one of the two poles of Eq. (6.14) is exactly the same as that in Eq. (6.12); another pole was at the origin.

Then, Eq. (6.14a) can be solved as follows:

$$\frac{1}{a([S_0]+\delta)[S_0])}d[S_0] = dt \tag{6.15a}$$

$$\frac{1}{a\delta}(\frac{1}{[S_0]} - \frac{1}{([S_0]+\delta)})d[S_0] = dt \tag{6.15b}$$

Integrating both sides simultaneously leads to:

$$\frac{1}{a\delta}(\ln[S_0] - \ln([S_0]+\delta)) = t + C_1 \tag{6.16a}$$

$$[S_0] = \frac{\delta e^{a\delta(t+C_1)}}{1-e^{a\delta(t+C_1)}} \Rightarrow C_1 = \frac{\ln\frac{[S_0]_0}{[S_0]_0+\delta}}{a\delta} \tag{6.16b}$$

where C_1 denotes the constant corresponding to the initial state of $[S_0]$, denoted as $[S_0]_0$.

Finally, after substituting Eq. (6.16b) into Eq. (6.14b), the following analytical solution of $[^1O_2]$ can be derived as:

$$\frac{d[^1O_2]}{dt} = c\frac{\delta e^{a\delta(t+C_1)}}{1-e^{a\delta(t+C_1)}} \tag{6.17a}$$

$$[^1O_2] = -\frac{c}{a}\ln(1-e^{a\delta(t+C_1)}) + C_2 \tag{6.17b}$$

$$C_2 = \frac{c}{a}\ln(1-e^{a\delta C_1}) \tag{6.17c}$$

where C_2 is a constant offset to ensure the initial value of $[^1O_2]$, i.e., the induced singlet oxygen by light at $t = 0$, to be zero. For the simplicity of the notations, denote $k_1 = -\frac{c}{a}$, $k_2 = e^{a\delta C_1}$, $k_3 = -\frac{a}{\phi}\delta$, $k_4 = C_2$. According to the definitions of a and c in Eq. (6.14), and C_1 in Eq. (6.16b), the parameters k_1, k_2, and k_3 actually take the following forms:

$$k_1 = \frac{f}{\sigma}, k_2 = e^{\ln\frac{[S_0]_0}{[S_0]_0+\delta}}, k_3 = \xi\sigma\delta. \tag{6.18}$$

Now, Eq. (6.17b) can be finally written as:

$$[^1O_2] = k_1\ln(1-k_2e^{-\phi k_3 t}) + k_4, \tag{6.19}$$

which is an analytical solution of $[^1O_2]$ in (6.1) in the case of abundant $^3O_2$. Note that although $k_1 \sim k_4$ are determined by the original parameters $[S_0]_0$, σ, δ, ξ, and f, fitting Eq. (6.19) actually does not require knowing or estimating these original ones. Therefore, only four parameters (k_1, k_2, k_3, and k_4) need to be estimated to fully know Eq. (6.19).

FUNCTIONAL RELATIONSHIP BETWEEN k_3 AND PHOTON ENERGY

In classical PDT models, the PDT dose is defined as the number of photons absorbed by the PS, and is related to the fluence rate and the photon energy. The kinetic PDT equations are henceforth parameterized by the fluence rate and photon energy of the stimulating light.

The energy of a photon is inversely proportional to the wavelength of the light, which is usually given in the unit of electron-volt (eV), i.e.:

$$E_p = \frac{hc}{\lambda} \tag{6.20}$$

where h and c denote the Planck's constant and the speed of light, respectively. Here, the unit of λ shall be converted from nanometers to microns. Photon energy will also be used to quantify a specific light source, as follows.

In Eq. (6.19), k_3 is the only parameter related to ξ, which is anti-proportional to the photon energy. Thus, the functional relationship between k_3 and E_p can be established as follows.

Based on our acknowledgments, the main PSs in *C. albicans* are flavin and porphyrin, whose peak absorptions are at 405nm and 460nm, respectively. Thus, the blue light of these two wavelengths can effectively inhibit the *C. albicans*. Moreover, k_3 is a parameter that reflects the singlet oxygen accumulation speed, which is correlated with the inhibition. Here, we use a piecewise function in Eq. (6.21) to describe this dependence of k_3 on the photon energy:

$$k_3 = \begin{cases} a_1 + b_1 \cdot sin(c_1 E_p + d_1), E_p \leq \overline{E} \\ a_2 + b_2 \cdot sin(c_2 E_p + d_2), E_p > \overline{E} \end{cases} \tag{6.21}$$

where $a_1, a_2, b_1, b_2, c_1, c_2, d_1, d_2$ are the parameters to be estimated; and $\overline{E}$ is demarcation point of the piecewise function that describes the peak near 415nm or 460nm.

MODELING $[^1O_2]$ AS A FUNCTION OF PHOTON ENERGY AND FLUENCE RATE

Now, by substituting Eq. (6.21) into Eq. (6.19) and noting that $\phi \cdot t = H_e$ (i.e., assuming a constant fluence rate), the model of the $^1O_2$ concentration y as a function of the photon energy and fluence can be finally derived in Eq. (6.22) on the top of next page.

$$y(H_e, E_p) = \begin{cases} k_1 \ln(1 - k_2 e^{-[a_1 + b_1 \cdot sin(c_1 E_p + d_1)] H_e}) + k_4, E_p \leq \overline{E} \\ k_1 \ln(1 - k_2 e^{-[a_2 + b_2 \cdot sin(c_2 E_p + d_2)] H_e}) + k_4, E_p > \overline{E} \end{cases} \tag{6.22}$$

Therefore, the $^1O_2$ concentration can be calculated using the photon energy and fluence rate of a treatment light source. This model can be applied in analyzing and calculating light dosages in clinical PDTs and ABLs. However, the fluence rate is found in clinical and *in-vivo* experiments to be highly subject to the optical properties of various tissues and the light wavelength [32]. In this study, we will especially focus on analyzing the light fluence rate in the superficial tissues of human body in what follows.

FLUENCE RATE FOR TREATING FUNGAL INFECTIONS IN CUTANEOUS TISSUES

For fungal infections in cutaneous tissues, i.e., on the surface of human body, external treatment light can be effectively applied to the skin. As a common practice in the literature [22, 32], a shallow layer of the skin tissue can be assumed as a semi-infinite medium with uniform optical properties, i.e., with the same absorption coefficient μ_a and the same reduced scattering coefficient μ_s'. On the other hand, for a planar light source that is applied externally, due to the backscattering effect from the tissue-air boundary, the

fluence rate beyond the boundary in the tissue can be even greater than that in the air [32]. That is

$$\frac{\phi}{\phi_{air}} = 1 + 2R_d, \tag{6.23}$$

where ϕ_{air} is the fluence rate of the incident light. R_d is the diffuse reflectance, and can be calculated as [32]:

$$R_d = \frac{a'}{2}\left(1 + e^{-\frac{4/3}{A}\sqrt{3(1-a')}}\right) \cdot e^{-\sqrt{3(1-a')}}, \tag{6.24}$$

where $a' = \mu_s'/(\mu_a + \mu_s')$ is the transport albedo; A is a constant that can be calculated based on the index of the refractions of the tissue and air at the boundary. For an air–tissue interface, $A = 2.9$ is a good approximation [22, 33]. For estimating the light fluence rate within the tissue, the following Lambart's relationship is used.

$$\phi(d) = \phi_{air} \cdot k \cdot e^{-\mu_{eff} \cdot d}, \tag{6.25}$$

where d is the depth below the tissue; μ_{eff} is the effective attenuation coefficient, which can be calculated as $\mu_{eff} = \sqrt{3 \cdot \mu_a(\mu_a + \mu_s')}$; and k is the backscatter coefficient, which is determined by [34] $k = 3 + 5.1R_d - 2e^{-9.7R_d}$.

The equations (6.23-6.25) suffice to characterize the light fluence rate in the cutaneous tissues below the air-skin boundary with a planar light source.

6.2.2. EXPERIMENTAL METHODS

CHEMICALS AND MATERIALS

The growth medium was tryptic soy broth (TSB), and was purchased from Hopebiol (Qingdao, China). The test assay of $^1O_2$ was Singlet Oxygen Sensor Green (SOSG) reagent (Invitrogen, Eugene, USA), which was dissolved in methanol (100 μg in 33 μl) to create a stock solution of 5 mM.

FUNGAL STRAINS

The *C. albicans* used in this work is the 3147 (IFO 1594) strain (ATCC, Manassas, VA, USA). The fungal strain was cultured overnight in TSB at 26^0C with shaking at 180 RPM. To maintain the concentration of the fungal suspension within the same range, the absorption of the suspension at 550nm was measured by a U-3900H spectrophotometer (Hitachi, Tokyo, Japan) before all the experiments. More specifically, the absorption levels measured by this equipment were always controlled in the range of 2.3 ~ 2.5, which corresponds to a fungal density of 10^7 CFU/ml.

LED LIGHT SOURCE

Four different types of LEDs with specified peak wavelengths at 385, 405, 415, and 450nm were selected to build four light sources. The spectral power density (SPD) curves for the LEDs, as measured by a Maya 2000 Pro spectrometer (Ocean Optics, Gloucester, MA, USA), are depicted in Fig. 6.1. In this figure, every SPD curve has been normalized with respect to its integral.

All the four light sources were made of a 4-by-4 LED array, arranged in a 1.5-by-1.5cm square as previously described [26]. Every light source was able to provide a round spot

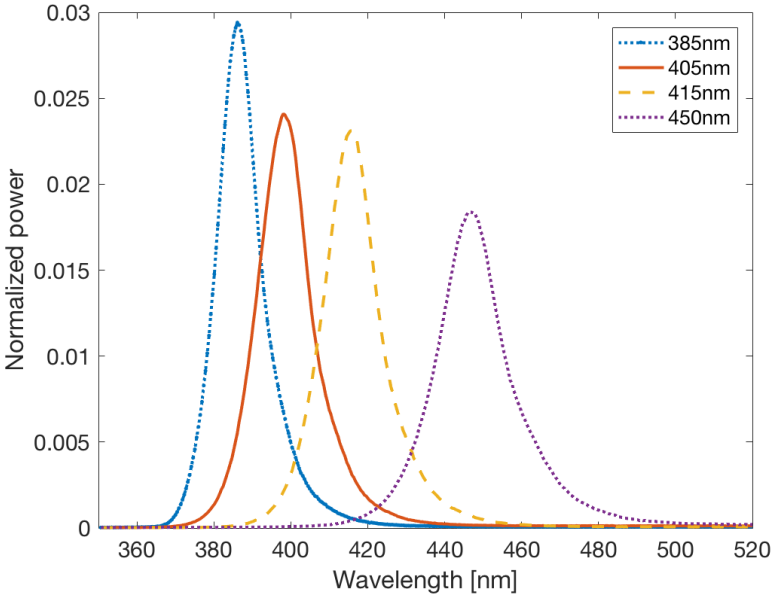


Figure 6.1: Normalized SPD curves of four LEDs.

with 15cm diameter at 10cm height above the sample. The target fluence rate in all the experiments was 50 mW/cm^2. The distribution of the irradiance in the target spot from every light source was measured using a PM100D power meter with a S120VC probe (Thorlabs, Newton, NJ, USA). According to the measurements, the average irradiance in the central 6cm-diameter circle was 49.39 mW/cm^2, with a relative variation of only 6.56 %. Therefore, the central light beam can be approximately regarded as being unidirectional; and the light source can be treated as a planar one. Thus, the measured irradiance and fluence rate are identical, both of which are numerically equal to the radiance in the direction of incidence of the photon beam [35].

SINGLET OXYGEN ASSAY

The SOSG reagent is highly selective for $^1O_2$. Unlike other available fluorescent and chemiluminescent singlet oxygen detection reagents, it does not exhibit any appreciable response to hydroxyl radical or superoxide. Before the assay, the fungal suspension was first washed by phosphate buffer saline (PBS) for three times. The SOSG stock solution was then added to it in a ratio of 1:1000. Next, the suspension was seeded into a 96-well plate and was irradiated by one of the four LED light sources. The cell suspension in one well was resuspended and placed in the dark at each sampling point of 0, 5, 10, 15, 20, 25, and 30 min.

Immediately after the light treatment, the singlet oxygen concentration levels of the cell suspension were measured as fluorescent levels using a VL0L0TD0 Varioskan LUX microplate reader (Thermo Fisher, Waltham, MA, USA), with the excitation and emission wavelengths set at 488nm and 525nm, respectively.

It is worth mentioning here that two potential factors that could affect the accuracy

of the assays were analyzed via preliminary experiments. Firstly, the temperature of the PBS solvent after being continuously irradiated for 30 min by the four light sources was measured, and was found to be lower than 37^oC in all the four cases. This indicated that the light irradiation could not lead to a temperature that was prohibitive to the *C. albicans*. Secondly, the absorption of the light by the PBS solvent was evaluated. The experimental measurements indicated that this absorption was indeed negligible, because of the colorless and transparent nature of the solvent. Therefore, the influence of both the light-induced heat and the light absorption by the PBS solvent on the light-induced $^1O_2$ was negligible.

STATISTICS

In the ROS assays, the raw data were processed to produce the mean and standard deviation for each treatment time interval. The significance of ROS levels and viability of cells were tested by the Student's t-test. The values of $P < 0.05$ were considered statistically significant.

6.3. RESULTS

6.3.1. SIMULATION RESULTS

SIMULATION AND ANALYSIS OF THE FIRST-PRINCIPLE MSO MODEL

By taking the MSO model (6.1) as the first-principle model, we simulated the ODEs of the model using Runge-Kutta-Fehlberg (RK45) method. The parameters were taken from a photofrin-mediated PDT model [21], i.e., $\sigma = 9 \times 10^{-5}$ μM^{-1}, $\xi = 3.7 \cdot 10^{-3}$ cm^2/mW/s, $\delta = 33$ μM, and $\beta = 11.9$ μM. The irradiance was set as ϕ=50 mW/cm^2. For the case with deficient oxygen concentration, $[^3O_2]_0$ and the supply rate g were set as in the normal condition as $[^3O_2]_0 = 50$ μM, $g = 0.5$ μM/s. On the other hand, in the case of rich oxygen concentration, their values were respectively set as $[^3O_2]_0 = 80$ μM and $g = 1$ μM/s, i.e., their maximum values mentioned in [21]. The irradiation time was set to 30 min with a step size of 5 min, which corresponded to the experimental time interval and total irradiation time.

In both cases, a sequence of time-series data with seven samples were generated, and were then taken as the equilibrium points to linearize the original nonlinear model (6.1). The root locus plots of these seven LTI systems are shown in Fig. 6.2. Two observations can be clearly made. First, the zero ranges from -5×10^{-5} to -2×10^{-5}, and cancels with the pole at the origin of all the seven LTI systems in both cases. Second, in the oxygen rich case, the root loci at the different points are condensed to a much smaller cluster, than those in the oxygen deficient case. This indicates that the LTI systems in the oxygen rich case are more identical to each other.

Then, the remaining two poles after the pole-zero cancelation were respectively calculated for both cases, and are shown in Fig. 6.3. Obviously, the variation of the poles in the oxygen deficient case is much larger than that in the oxygen rich case. Specifically, the magnitude of the non-dominant pole drops from 50.98 % to 12.85 %, when the oxygen concentration increases; while the drop in the dominant pole is even more remarkable, i.e., from 59.37 % to 9.14 %.

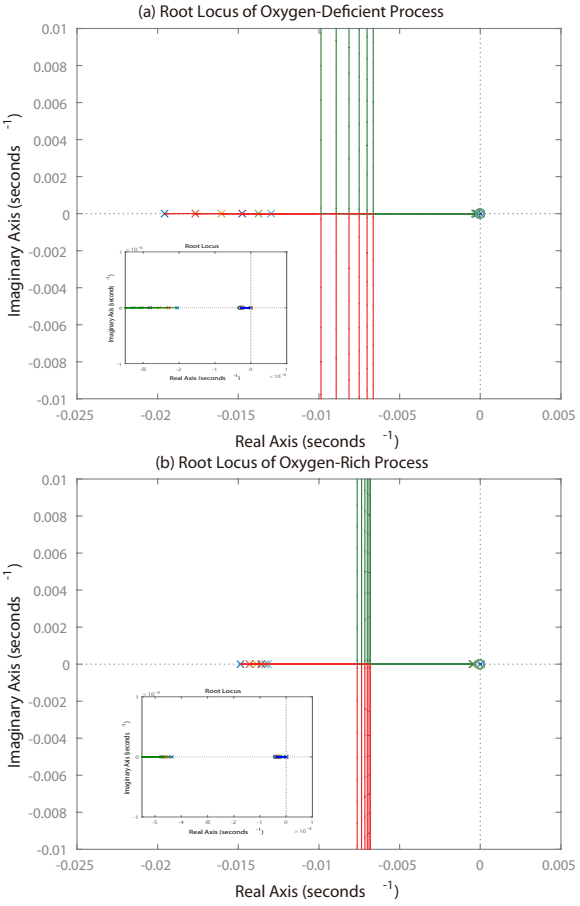


Figure 6.2: The root locus plots of the seven linearized MSO processes at different equilibrium points: (a) $^3O_2$ deficient case, (b) $^3O_2$ rich case. The scales of the corresponding axes in the two figures are the same. The inserts in the figures show the cancelation of the pole $s = 0$ with the zero.

VALIDATION OF THE LINEARIZED MODEL (6.6) TO MODEL (6.1) BY SIMULATIONS

In this section, the linearized model (6.6) was used to simulate $[^1O_2]$ in an oxygen-rich environment. Because this system can be regarded as a linear system with an abundant oxygen supply, $\boldsymbol{A}$ and $\boldsymbol{B}$ were fixed to their mean value in Eq. (6.6). The fitting result is plotted in Fig. 6.4(a) with a 8.5366 average root mean square error (RMSE).

VALIDATION OF THE CLOSED-FORM SOLUTION (6.19) TO MODEL (6.1) BY SIMULATIONS

In this section, the closed-form solution of the first-principle model was verified under different oxygen supplies. We first consider the situation of an ample oxygen supply. The oxygen-concentration related parameters were set as $g = 0.8\ \mu$M/s, $[^3O_2] = 50\ \mu$M. The simulated results between the closed-form solution (6.19) and model (6.1) are shown in Fig. 6.4(b). The estimated parameters are shown in Tab. 6.1. The RMSE of the fitted

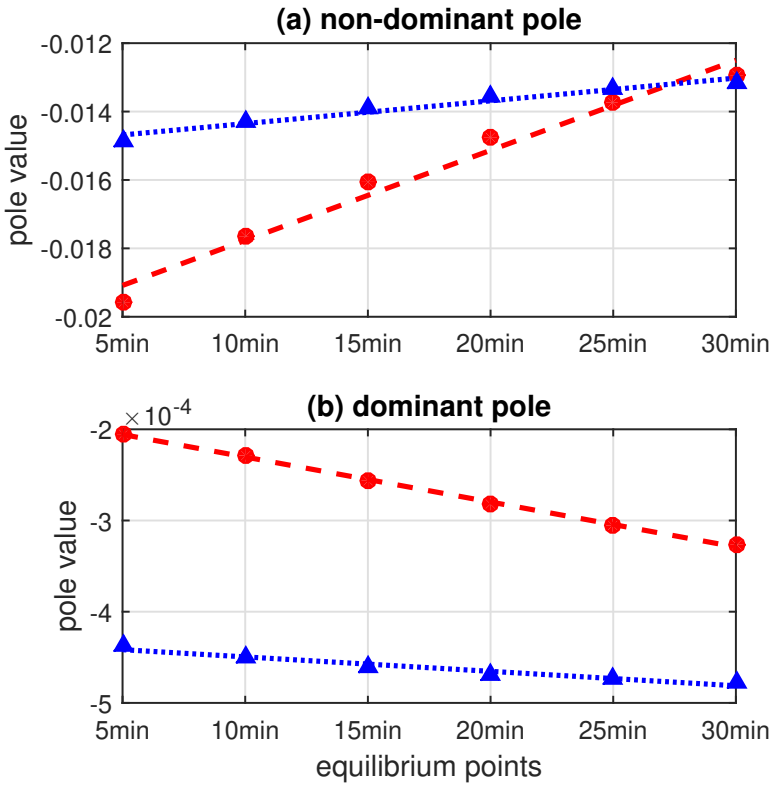


Figure 6.3: The pole positions of the linearized MSO process at different equilibrium points: (a) non-dominant pole, (b) dominant pole. Cycles: ${}^3O_2$ deficient case. Triangles: ${}^3O_2$ rich case.

model (6.19) is only 1.2087, which obviously demonstrates its accuracy.

Table 6.1: Estimated parameters of Eq. (6.19).

k_1	k_2	k_3	k_4	RMSE
6.289e+03	0.0920	7.168e-06	609.553	1.2087

The entire simplification is based on the approximation of an abundant oxygen supply. Thus, we verified the closed-form solution Eq. (6.19) with different oxygen supply rates g. The Mean relative absolute errors (MRAE), defined as follows, of Eq. (6.19) compared to the first principle model with different oxygen supply rates are depicted in Fig. 6.5.

$$MRAE = \frac{1}{T}\sum_{t=1}^{T}\frac{\left|\widehat{[{}^1O_2]}_t - [{}^1O_2]_t\right|}{[{}^1O_2]_t} \tag{6.26}$$

The perfect fit of the closed-form solution Eq. (6.19) to the simulated data motivated us to fit the experimental data to this model.

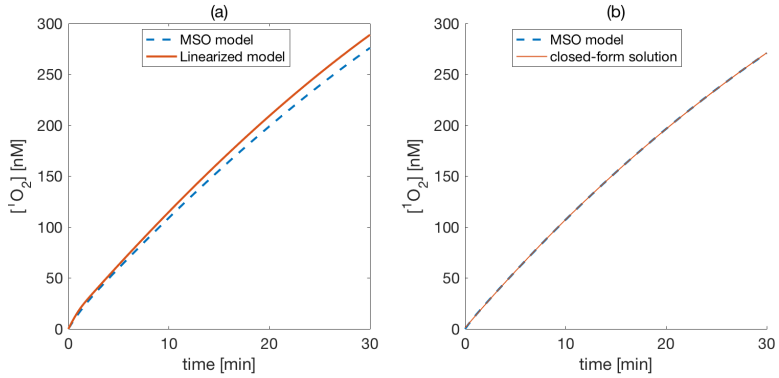


Figure 6.4: The accuracy of the (a) linearized model and (b) closed-form solution comparing to the first-principle model.

6.3.2. EXPERIMENTAL RESULTS

MEASUREMENTS OF THE SINGLET OXYGEN CONCENTRATIONS IN *C. albicans* AT ABL IRRADIATION

The time sequences of the singlet oxygen concentrations in the *C. albicans* were measured from the experiments using all four different light sources, as shown in Fig. 6.6. The singlet oxygen concentration was significantly increased after 10 min irradiation of 385nm wavelength (P<0.05). In other cases, a significant increase in $[^1O_2]$ was observed within 5 min of irradiation (P<0.05). The significance of the singlet oxygen concentration in fungi before and after the light treatment was tested by the Student's t-test.

FITTING THE PARAMETERS OF EQ. (6.19) TO THE EXPERIMENTAL DATA

Due to the aforementioned challenges in parameterizing the model of the induced $[^1O_2]$ in vitro, a closed-form solution of the MSO process was developed. One needs to estimate four parameters (k_1, k_2, k_3, and k_4) to use this model. More specifically, k_1 and k_2 rely on the type of pathogen and its endogenous PS; k_4 is an offset to the initial concentration of singlet oxygen, whereas k_3 relies on the type of the pathogen and the PS, and also on the photochemical parameters that are related to the photon energy.

Motivated by the definition of these parameters, we first fix the values of k_1, k_2, and k_4, and then built a functional relationship between k_3 and photon energy to investigate the light-induced $^1O_2$ of *C. albicans* by different light wavelengths.

First, a standard least-squares (LS) algorithm was used to coarsely tune the values of k_1, k_2, k_3, and k_4. The values of these parameters fitted to the experimental data with different wavelengths are listed in Tab. 6.2. It is found that k_1 and k_2 were almost constant regardless of the wavelengths.

The MRAE errors for the values of k_1 and k_2 are shown in Fig. 6.7. It can be seen that in all the four cases, the optimum of the parameter pair (k_1,k_2) lies on the flat bottom of a narrow valley roughly within the range of $10 < k_1 < 10^4$ and $0.1 < k_2 < 1$. Thus, we set $k_1 = 280$ and $k_2 = 0.9$ as constant for all cases.

Second, with fixed k_1, k_2, again, we used the LS algorithm to fine-tune k_3 and k_4. The results are listed in Tab. 6.3 and plotted in Fig. 6.8. k_4 was found to be approximately

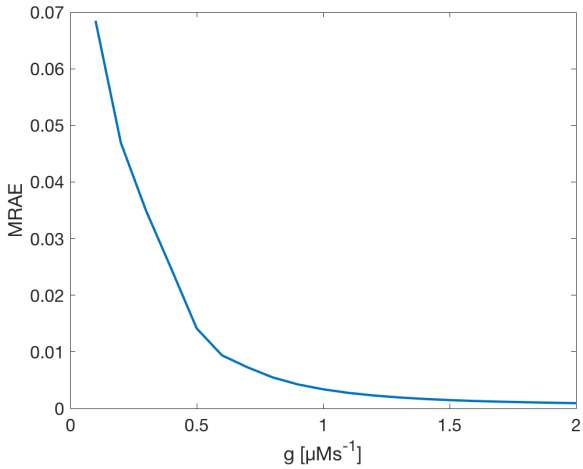


Figure 6.5: The fitting accuracy of the closed-formed solution (6.19) as compared to the first-principle model (6.1) in different oxygen supply rate g.

Table 6.2: Coarsely tuned values of the parameters of Eq. (6.19) by least squares

Wavelength [nm]	k_1	k_2	k_3	k_4	**RMSE**
385	305.8660	0.8971	1.400e-5	716.1151	5.6245
405	257.9358	0.9399	8.000e-6	749.1535	4.3801
415	281.9836	0.9855	2.031e-06	1197.8264	1.5716
450	318.8462	0.7893	1.8e-5	517.9739	1.4840

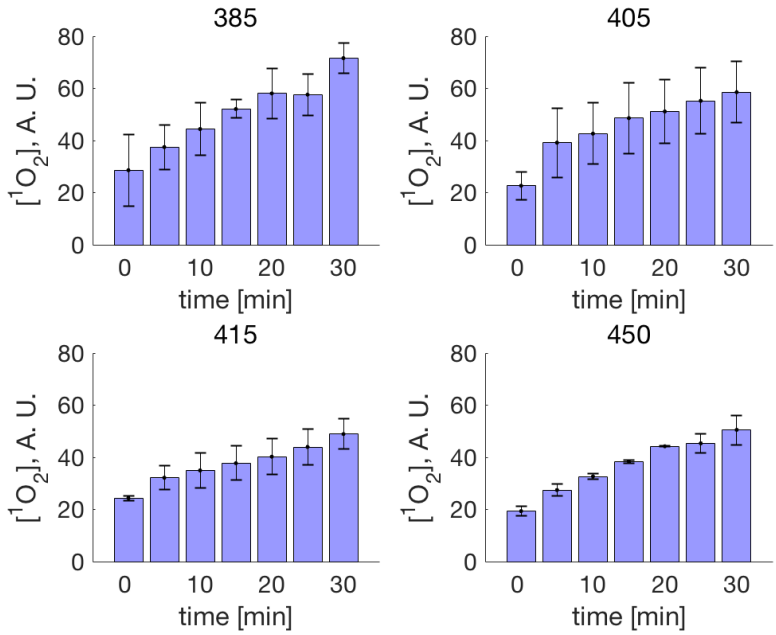


Figure 6.6: The singlet oxygen concentration in *C. albicans* induced by different light wavelength.

Table 6.3: Fine tuned values of k3 and k4 in Eq. (6.19)

Wavelength [nm]	k_3	k_4	**RMSE**
385	1.1240e-05	674.6588	2.4744
405	9.0556e-06	673.8206	3.4526
415	6.0888e-06	671.0575	1.2636
450	8.5290e-06	666.1381	1.4852

constant in the results and can therefore also be fixed at $k_4 = 670$. In contrast, k_3 was sharply altered in different experiments.

In Eq. (6.19), k_3 determines the speed of generating $^1O_2$ in the photochemical reaction. The larger the value, the faster the accumulation of the cytotoxic $^1O_2$.

RELATIONSHIP BETWEEN k_3 AND PHOTON ENERGY

To fit and validate the relationship between k_3 and photon energy, we split the experimental data into two groups, i.e., with those from the 385nm, 405nm, and 450nm light experiments for fitting, and with the 415nm experimental data for validation.

In Eq. (6.21), there are totally nine unknown parameters to be estimated, i.e., a_1, a_2, b_1, b_2, c_1, c_2, d_1, d_2, and E. However, there are only three data points, which are not enough to uniquely determine all these nine parameters by solving a single LS problem. To fit them, c_1 and c_2 were first determined by estimating the periods of the sinusoidal functions. More specifically, c_1 was greater than c_2 which corresponds to the PS of the *C. albicans* absorption spectra. Then, b_1 and b_2 were secondly estimated by the heights

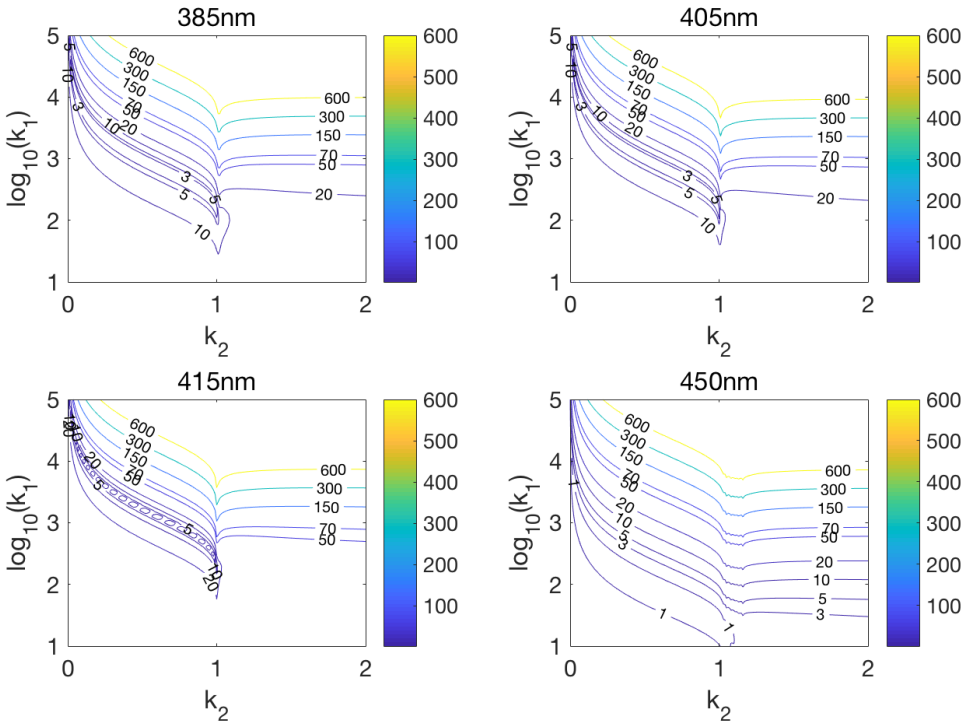


Figure 6.7: The fitting errors with different values of k_1 and k_2, where the numbers on the color bars and the contours are the MRAE errors.

of the peaks. For the same reason, b_1 is smaller than b_2. This then reduces the number of unknown parameters to five, where a_1, a_2, d_1 and d_2 can now be estimated by a nonlinear LS algorithm with reasonable accuracy; and $\overline{E}$ is finally determined by the intersection of piecewise functions. As a result, the value of k_3 in the fitted model at 415nm turned out to be 5.855×10^{-6} wihch is plotted on Fig. 6.9. As a validation, the fitted k_3 by the experimental data was 6.089×10^{-6} as reported in Tab. 6.3. The difference between the model output and the experimentally fitted k_3 at this validation point was about 4%, which is acceptable.

MODELING SINGLET OXYGEN AS A FUNCTION OF PHOTON ENERGY AND FLUENCE RATE

With the parameters in Tab. 6.4 on page 9, the model of $[^1O_2]$ as a function of the stimulating photon energy and fluence rate, i.e., Eq. (6.22), was finally established, with the validation using the 415nm (E_p=2.99 eV) experimental data. The simulation results are plotted in Fig. 6.10 where the validating data are shown in squares. The RMSE error of the predicted $[^1O_2]$ from the model compared to the validating experimental measurements was only 1.251. Therefore, the accuracy of this model at the validation point is also acceptable. However, the limitation of this model is that it is accurate only in an oxygen-rich environment. According to our simulation studies, a rule of thumb to determine the richness of oxygen supply is that the supply rate shall be higher than 0.4 μMs^{-1}

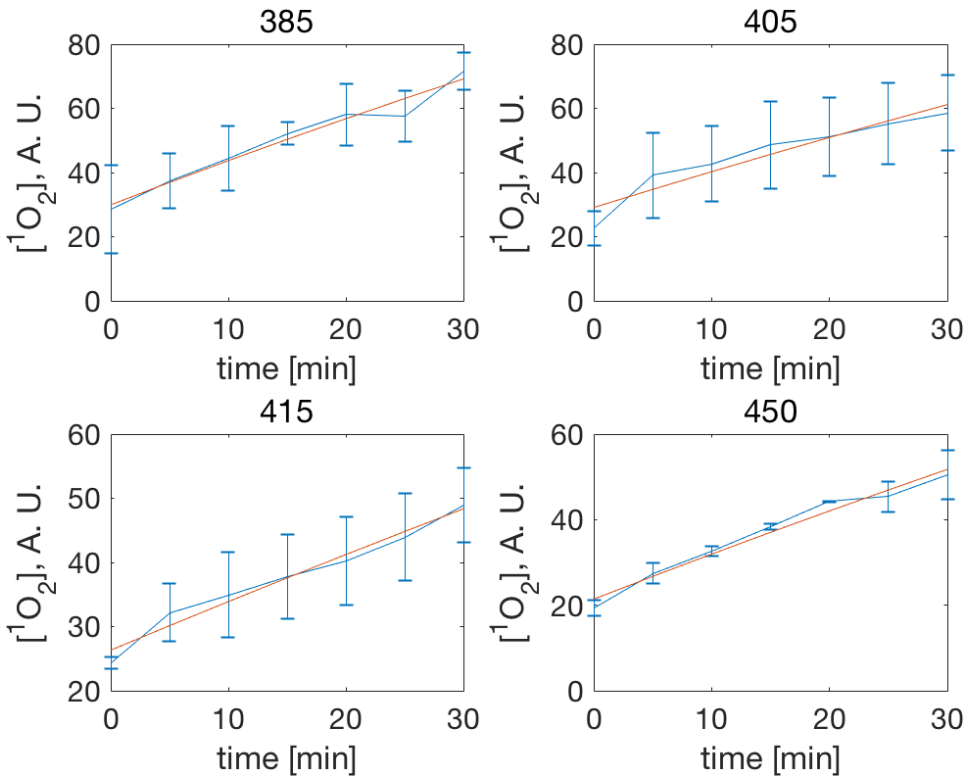


Figure 6.8: Fitted curve to the measured [$^1O_2$] from different wavelength light irradiation experiment.

for *in vivo* or clinical experiments.

Table 6.4: Estimation of the parameters of Eq. (6.22)

a_1	a_2	b_1	b_2	c_1	c_2	d_1	d_2	k_1	k_2	k_4
6.655e-06	6.537e-6	8e-6	2e-6	14	19	0.999	0.960	280	0.9	670

From the simulation results of (6.22), the following observations can be further made. First, when the fluence rate is low, i.e., H_e <5 J/cm^2, $^1O_2$ is insufficiently induced by any photon energy in the range of 2.6 ~ 3.2 eV. This indicates that sufficiently large radiometric energy of the stimulating light is required to trigger sufficient singlet oxygen generation, no matter how large the photon energy is. Second, as the fluence rate accumulates to a certain level, i.e., H_e >15 J /cm^2 , the photon energy becomes the main factor to determine [$^1O_2$]. For the fungi studied in this study, *C. albicans*, the photon energy in the range of either 2.7 ~ 2.8 eV or 3.0 ~ 3.2eV (388 ~ 413nm or 443 ~ 459nm in wavelength) can more effectively stimulate singlet oxygen generation, which coincides with the Sorét bands of common types of porphyrin and flavin.

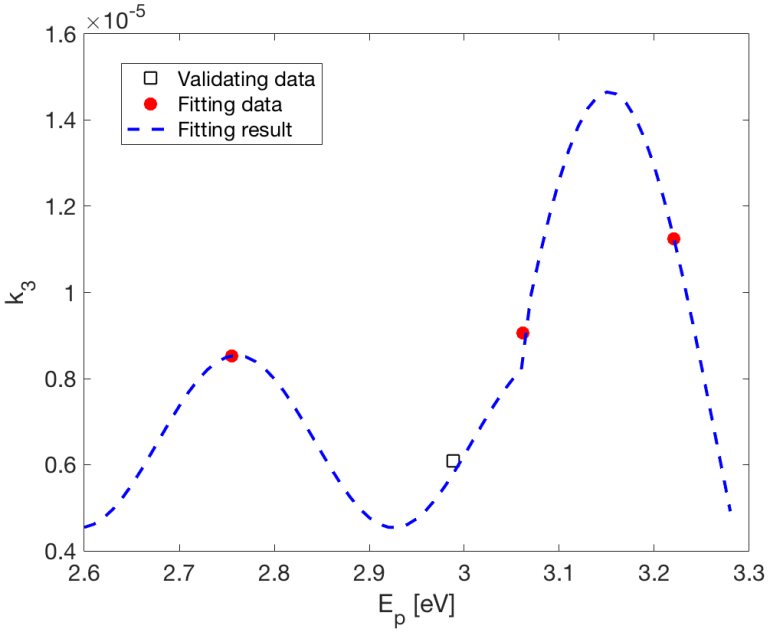


Figure 6.9: Fitted model of k_3 as a function of the photon energy.

QUANTIFYING THE FLUENCE RATE FOR TREATING FUNGAL INFECTIONS IN CUTANEOUS TISSUES

To study the clinical effect of ABL on treating superficial infections, knowing the fluence rate delivered to the tissues beneath the surface is important. To this end, we first collected the optical properties of healthy human epidermis from the literature, including the absorption coefficient μ_a and the reduced scattering coefficient μ_s'. Due to the limited measurements reported so far, it was still not possible to find all the values exactly corresponding to the four wavelengths. Fortunately, in the studied range of wavelength, the available values were found to be μ_a=1.5 cm^{-1}, μ_s'=10 cm^{-1} at 400nm, and μ_a=1 cm^{-1}, μ_s'=9 cm^{-1} at 450nm [36]. By respectively applying these parameters in Eqs. (6.23-6.25), we simulated the light fluence rate up to 1cm depth beneath the epidermis. The result is shown in Fig. 6.11(a). As can be seen, the light fluence rate increases sharply at the air-tissue boundary because of the backscattering from the boundary; and then exponentially decreases with the increasing depth. More specifically, the fluence rate of the 400nm light decreases faster than that of the 450nm. For instance, the fluence rate of 450nm at 0.2cm depth is 5.5 times lager than that of 400nm. That is to say, the 450nm can penetrate deeper than 400nm.

Furthermore, the spatial distribution of the induced $^1O_2$ concentration was simulated using Eq. (6.22), with a total fluence rate of 90 J/cm^2. The simulation result is shown in Fig. 6.11(b), where the $^1O_2$ concentration exponentially decreases with the increasing depth.

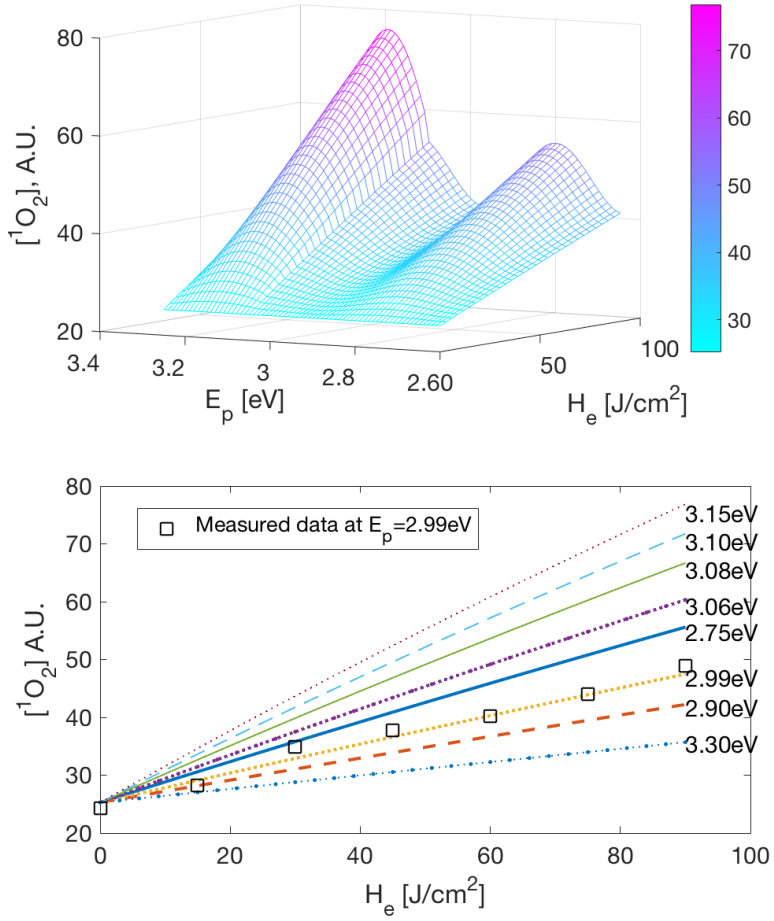


Figure 6.10: $[^1O_2]$ in response to various photon energy and fluence rate, where the numbers on the color bar indicate the values of $[^1O_2]$.

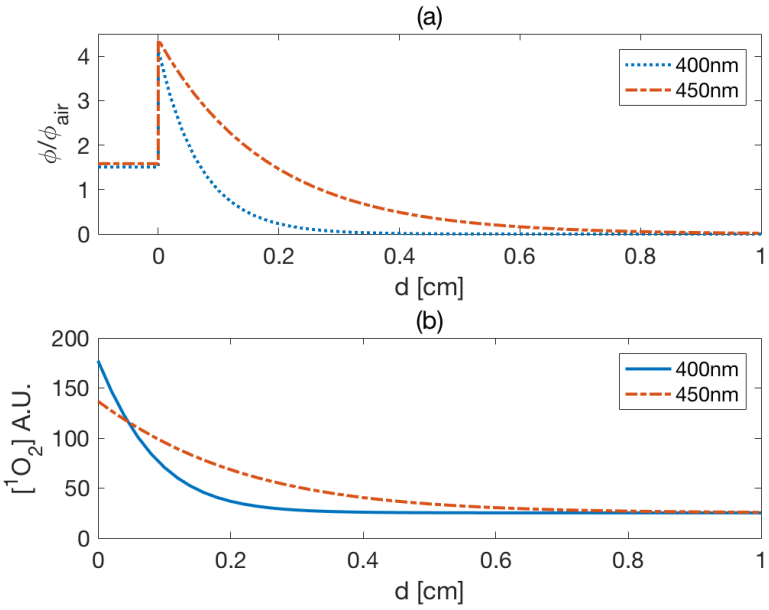


Figure 6.11: The spatial distributions of (a) fluence rate and (b) $^1O_2$ concentration up to 1cm depth within human epidermis.

6.4. DISCUSSION

6.4.1. CONDITIONS AND ACCURACY OF THE LINEARIZED MODEL (6.6)

According to both the root loci in Fig. 6.2 and the pole positions in Fig. 6.3, the nonlinearity of the MSO process (6.1) is greatly reduced in an oxygen rich environment. The process can thus be approximately treated as linear. It is worth mentioning that it is much easier for such a linear model to be obtained from the first-principle model or identified from time-series data. On the other hand, the dominant poles λ_2 of the oxygen rich case are faster than their counterparts in the oxygen deficient case. This also indicates that the response of $^1O_2$ to light stimulation is faster, when the main ingredient $^3O_2$ is abundant.

More specifically, as analyzed in Eq. (6.12), when the concentration of triplet oxygen $[^3O_2]$ is far greater than β, $[S_0]$ will become the only time-varying parameter that can change the pole. In the simulation, although $[^3O_2]$ is not infinite, it is still approximately five times more than β with a normal oxygen content environment; this ratio increases to ten times greater in an oxygen-rich environment. Thus, $[S_0]$ is the main factor that determines the pole positions in an oxygen-rich environment. Moreover, since σ is far less than the other parameters in Eq. (6.1), i.e., $\sigma = 9{\cdot}10^{-5}\ \mu\text{M}^{-1}$, $\xi = 3.7{\cdot}10^{-3}\ \text{cm}^2/\text{mW/s}$, $\delta = 33\ \mu\text{M}$, $\beta = 11.9\ \mu\text{M}$, the variation of $[S_0]$ is relatively lower than that of the other components. In the simulation, $\frac{d[S_0]}{dt}$ was only about 1% of $\frac{d[^3O_2]}{dt}$ and $\frac{d[^1O_2]}{dt}$ in the entire process. Therefore, there was little change in the pole positions over time in the oxygen-rich environment, and thus, the dynamics could be approximated to those of a linear

system. This provides the insight into the dynamic behavior of the MSO process, in the sense that the MSO process is approximately linear in the presence of ample oxygen concentration.

For a linear system, a linearized model (6.6) was used to simulate the changes in $^1O_2$. However, the fitting results were suboptimal as compared to using the nonlinear model (6.19). The reason for this can be the nonlinearity of the MSO model. As we previously discussed, the MSO model can be approximated to linear in an oxygen-rich environment. However, the poles are slowly changing over time, and strictly speaking, it remains a nonlinear model. This causes error in the linearized model, which assumes the fixed system poles and zeros.

6.4.2. CONDITIONS AND ACCURACY OF THE MODEL EQ. (6.19)

The closed-form solution is based on the condition of abundant oxygen supply. The concentration of $^3O_2$ was assumed to be infinite in the derivation of Eq. (6.19). Although $[^3O_2]$ cannot be infinite in reality, when the value of $[^3O_2]$ is sufficiently large, the accuracy of the closed-form solution is acceptable. In Fig. 6.5, we validated this solution under different oxygen supply rates. More specifically, when the oxygen supply rate $g > 0.4\mu Ms^{-1}$, the MRAE of Eq. (6.19) is always higher than 98%. When $g = 0.8\mu Ms^{-1}$, shown in Fig. 6.4(b), the accuracy of the closed-form solution can be up to 99%. In reality, the oxygen supply rate in human tissues is in the range of $0.1 - 1\mu Ms^{-1}$ [21]. Therefore, in most cases, this simplified closed-form model can still give accurate enough results.

It shall be mentioned that knowing the level of oxygen during PDTs is also important to predict the PDT efficacy [37], since the cytotoxic effect of PDTs depends in part on the availability of molecular oxygen in the target tissue. In this aspect, it was reported that high fluence rates could lead to a reduced photodynamic effect due to the oxygen depletion [38]. Therefore, for the accuracy of our model, it shall also be recommended to avoid apply high fluence rates in clinical treatment. On the other hand, applying a pulsed light with high amplitude can also be a way out. To determine the optimum timing of the duty cycles of the light pulses, a mathematical model can be used to analyze their effects on intercapillary tissue during PDTs [39].

On the other hand, it shall also be noted that the accuracy of the closed-form MSO model is dependent on the oxygen supply rate g. According to our simulation studies, a rule of thumb to determine the richness of oxygen supply is that the supply rate shall be higher than $0.4\mu Ms^{-1}$. In reality, g also depends on blood flow rate; and $g = 0.7\mu Ms^{-1}$ is a well-accepted value in PDTs [22, 40], which clearly meets the condition of an accurate closed-form MSO model. Besides, in treating superficial infection, due to the direct exposure of the treatment area to the air, the oxygen supply can be even higher. Therefore, the oxygen-rich environment assumption can be satisfied in both cases, and does not impose any limitation to apply the model.

6.4.3. CLINICAL RELEVANCE OF THE CLOSED-FORM MODEL (6.19)

In traditional PDTs, reduced tumor sizes and cure index are usually considered as their outcomes, which have been found to be highly correlated with the concentration of singlet oxygen [22, 24]. Therefore, the first principle model (6.1) is widely applied clinically

to determine the specified light dosage with a target $[^1O_2]$. Since the closed-form MSO model (6.19) is directly derived from Eq. (6.1), it can be equivalently used to evaluate the PDT outcomes, under the aforementioned condition of ample oxygen supply.

It shall be emphasized that although Eq. (6.19) was especially tested in treating fungal infections, this model can actually be applied to all cases of antimicrobial light therapies and PDTs. In any case, one only needs to estimate the four parameters (k_1, k_2, k_3 and k_4) to establish the model. To this end, a data-driven method was used in Sec. 6.3.2 to estimate them. On the other hand, these parameters can also be calculated with the knowledge of the PS dependent parameters, i.e., $[S_0]_0$, σ, δ, ξ and f. These parameters can be obtained either from the relevant literature or by conducting experiments following standard procedures [23]. Once $k_1 \sim k_4$ are determined, $[^1O_2]$ can be calculated with a specified fluence by Eq. (6.19). The $^1O_2$ concentration can then be used to predict the PDT outcomes. In traditional PDTs, the relationship between the $^1O_2$ concentration and treatment outcome can be found in many literatures. For instance, when $[^1O_2] > 1.1mM$, the radiation-induced fibrosarcoma tumors exhibited a complete response after 14 days PDT treatment [40]. For the more susceptible bacteria, it was reported that *S. typhimurium* and *E. coli* were inactivated by 99% with $[^1O_2]$ accumulated to $74\mu M$ [41]. As a summary, the procedures of establishing the model (6.19) for a potential clinical application are illustrated in Fig. 6.12.

6.5. CONCLUSION

In this work, the nonlinearity of the MSO process was examined by the poles of its linearized models at different equilibrium points. It was found that this system is linear in an oxygen-rich environment. Thus, the MSO model can be reduced to a linear dynamic model. Although this linearized model provides additional insights into the dynamic behavior of the MSO process, it can also lead to approximation errors in the prediction of long-time trajectories. Hence, an closed-form exact solution of MSO model was derived. Furthermore, based on this analytical solution, a general model was developed as a function of the photon energy and light fluence rate.

We also measured the singlet oxygen concentration of *C. albicans* up to 30 min with the irradiation of 385, 405, 415, and 450nm light. The fitting results of our model are in agreement with the experimental data, indicating the effectiveness of the proposed model and parameter estimation methods. We discovered that the photon energy within the range of 2.7 ~ 2.8 eV or 3.0 ~ 3.2 eV (388 ~ 413nm or 443 ~ 459nm in wavelength) is more effective in generating singlet oxygen in the fungi studied in this work, and is hence more effective in treating the infections caused by this type of fungi.

A potential extension of the current work is to model and compare the light induced singlet oxygen in other types of infectious fungal strains based on the developed analysis and modeling methods.

REFERENCES

[1] T. J. Dougherty *et al.*, *Photodynamic therapy,* J. Natl. Cancer Inst. **90**, 889 (1998).

[2] T. Maisch, *A new strategy to destroy antibiotic resistant microorganisms: antimicro-*

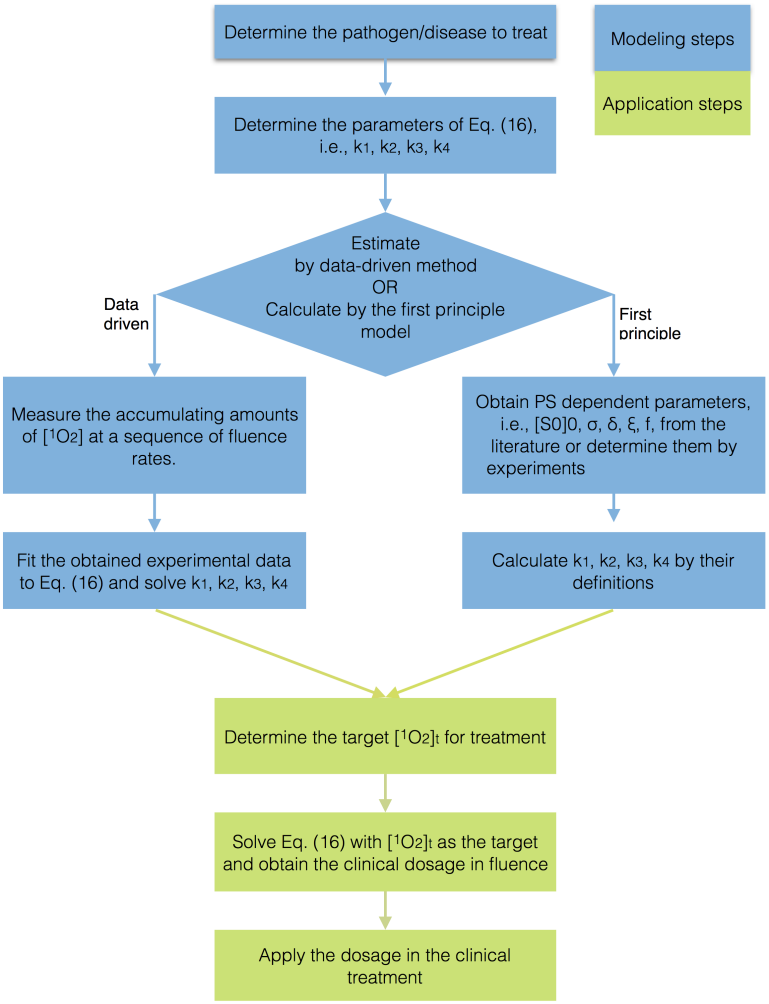


Figure 6.12: Procedures in applying the closed-form algebraic model (6.19)

bial photodynamic treatment, Mini Rev. Med. Chem. **9**, 974 (2009).

[3] T. Dai *et al., Blue light for infectious diseases: Propionibacterium acnes, helicobacter pylori, and beyond?* Drug Resist. Updat. **15**, 223 (2012).

[4] J. S. Guffey *et al., Inactivation of mycobacterium smegmatis following exposure to 405-nanometer light from a supraluminous diode array.* Wounds **25**, 131 (2013).

[5] M. Dring, *Photocontrol of development in algae,* Annual Review of Plant Physiology and Plant Molecular Biology **39**, 157 (1988).

[6] Y. Zhang *et al., Antimicrobial blue light inactivation of candida albicans: In vitro and in vivo studies,* Virulence **7** (2016).

[7] F. Halstead *et al., Violet-blue light arrays at 405 nanometers exert enhanced antimicrobial activity for photodisinfection of monomicrobial nosocomial biofilms,* Appl. Environ. Microbiol. **85** (2019).

[8] R. Yin *et al., Light based anti-infectives: ultraviolet c irradiation, photodynamic therapy, blue light, and beyond,* Curr. Opin. Pharmacol. **13**, 731 (2013).

[9] R. R. Allison and K. Moghissi, *Photodynamic therapy (PDT): PDT mechanisms,* Clin. Endosc. **46**, 24 (2013).

[10] M. Hayyan *et al., Superoxide ion: generation and chemical implications,* Chem. Rev. **116**, 3029 (2016).

[11] G. Pannov *et al., The role of iron in N_2O decomposition on ZSM-5 zeolite and reactivity of the surface oxygen formed,* J. Mol. Catal. **61**, 85 (1990).

[12] J.-P. Pouget *et al., DNA damage induced in cells by γ and UVA radiation as measured by HPLC/GC- MS and HPLC- EC and comet assay,* Chem. Res. Toxicol. **13**, 541 (2000).

[13] J. F. Turrens, *Mitochondrial formation of reactive oxygen species,* J. Physiol. **552**, 335 (2003).

[14] T. C. Zhu and J. C. Finlay, *The role of photodynamic therapy (PDT) physics,* Med. Phys. **35**, 3127 (2008).

[15] A. E. Profio and D. R. Doiron, *Dosimetry considerations in phototherapy,* Med. Phys. **8**, 190 (1981).

[16] J. Dysart *et al., Sci-YIS Fri-03: Calculation of singlet oxygen dose from photosensitizer photobleaching during mthpc or photofrin photodynamic therapy in vitro,* Med. Phys. **32**, 2419 (2005).

[17] I. Georgakoudi *et al., The mechanism of photofrin photobleaching and its consequences for photodynamic dosimetry,* Photochem. Photobiol. **65**, 135 (1997).

[18] X.-H. Hu, , *et al., Modeling of a type II photofrin-mediated photodynamic therapy process in a heterogeneous tissue phantom,* Photochem. Photobiol. **81**, 1460 (2005).

[19] T. C. Zhu *et al.*, *Macroscopic modeling of the singlet oxygen production during PDT,* in *Optical Methods for Tumor Treatment and Detection: Mechanisms and Techniques in Photodynamic Therapy XVI,* Vol. 6427 (International Society for Optics and Photonics, 2007) p. 642708.

[20] K. W. Beeson *et al.*, *Validation of combined monte carlo and photokinetic simulations for the outcome correlation analysis of benzoporphyrin derivative-mediated photodynamic therapy on mice,* J. Biomed. Opt. **24**, 035006 (2019).

[21] G. Kareliotis, S. Liossi, and M. Makropoulou, *Assessment of singlet oxygen dosimetry concepts in photodynamic therapy through computational modeling,* Photodiagnosis Photodyn. Ther. **21**, 224 (2018).

[22] K. K.-H. Wang *et al.*, *Explicit dosimetry for photodynamic therapy: macroscopic singlet oxygen modeling,* J. Biophotonics **3**, 304 (2010).

[23] M. M. Kim *et al.*, *On the in-vivo photochemical rate parameters for PDT reactive oxygen species modeling,* Phys. Med. Biol. **62**, R1 (2017).

[24] T. Sheng *et al.*, *Reactive oxygen species explicit dosimetry to predict local tumor growth for photofrin-mediated photodynamic therapy,* Biomed. Opt. Express **11**, 4586 (2020).

[25] Y. Gu *et al.*, *Nonlinear modeling of cortical responses to mechanical wrist perturbations using the narmax method,* IEEE Trans. Biomed. Eng. **68**, 948 (2020).

[26] J. Dong and T. Wang, *Data driven modeling of the reactive oxygen species stimulated by photon energy in light therapies,* IEEE Access **8**, 18196 (2020).

[27] T. Wang *et al.*, *Analyzing efficacy and safety of anti-fungal blue light therapy via kernel-based modeling the reactive oxygen species induced by light,* IEEE Trans. Biomed. Eng. **In Press** (2022).

[28] R. A. Calderone and W. A. Fonzi, *Virulence factors of candida albicans,* Trends Microbiol. **9**, 327 (2001).

[29] T. Arendorf and D. Walker, *The prevalence and intra-oral distribution of candida albicans in man,* Arch. Oral Biol. **25**, 1 (1980).

[30] J. D. Sobel, *Vaginitis,* N. Engl. J. Med. **337**, 1896 (1997).

[31] R. C. Dorf and R. H. Bishop, *Modern control systems* (Pearson Prentice Hall, 2008).

[32] A. Dimofte *et al.*, *In vivo light dosimetry for motexafin lutetium-mediated pdt of recurrent breast cancer,* Lasers Surg. Med. **31**, 305 (2002).

[33] J. L. Sandell and T. C. Zhu, *A review of in-vivo optical properties of human tissues and its impact on pdt,* J. Biophotonics **4**, 773 (2011).

[34] S. L. Jacques, *Light distributions from point, line and plane sources for photochemical reactions and fluorescence in turbid biological tissues,* Photochem. Photobiol. **67**, 23 (1998).

[35] T. C. Zhu *et al.*, *Absolute calibration of optical power for PDT: Report of AAPM TG140,* Med. Phys. **40**, 081501 (2013).

[36] E. V. Salomatina *et al.*, *Optical properties of normal and cancerous human skin in the visible and near-infrared spectral range,* J. Biomed. Opt. **11**, 064026 (2006).

[37] J. Moan and S. Sommer, *Oxygen dependence of the photosensitizing effect of hematoporphyrin derivative in nhik 3025 cells,* Cancer Res. **45**, 1608 (1985).

[38] T. H. Foster *et al.*, *Oxygen consumption and diffusion effects in photodynamic therapy,* Radiat. Res. **126**, 296 (1991).

[39] J. P. Henning *et al.*, *A transient mathematical model of oxygen depletion during photodynamic therapy,* Radiat. Res. **142**, 221 (1995).

[40] H. Qiu *et al.*, *A comparison of dose metrics to predict local tumor control for photofrin-mediated photodynamic therapy,* Photochem. Photobiol. **93**, 1115 (2017).

[41] T. A. Dahl *et al.*, *Comparison of killing of gram-negative and gram-positive bacteria by pure singlet oxygen,* J. Bacteriol. **171**, 2188 (1989).

7

CONCLUSION

7.1. SUMMARY OF RESULTS

In this thesis, we focused on investigating the safety and efficacy of ABL therapies. Aiming on the dosage optimization in ABL, several dynamic nonlinear models were developed. Here are the main significance and founding of this thesis:

- First, the safety and efficacy of ABL was investigated in Chapter 2 by *in vivo* experiments. In that work, the inactivating effects of ABL on both the pathogens and human host cells. A dynamic viability models and functional relationship of these two types of cells were built based on the experimental data. The results indicate that, in vitro, the 415nm light source has a more effective anti-fungal function with less damage to the human host cells than 450nm and 405nm wavelength.

- Second, we measured the ROS concentration on the aforementioned two types of cell during the ABL irradiation in . Furthermore, a first order linear integrator has been developed to mathematically describe the induced ROS in fungi, as a function of the photon energy and fluence of the stimulating light in Chapter 3. The fitting results agree well with the main trends of the experimental data at long time scales, e.g. from tens of minutes to a few hours.

  This linear integrator model was applied on simulating the nonlinear ROS accumulation process. Obviously, the theoretical error can not be eliminated. Therefore, more nonlinear models were developed on solving the problem.

- Third, new modeling scheme was proposed in Chapter 4 using a kernel-based NARX structure, whose kernels were selected by a newly developed algorithm based on PSO optimization. This NARX model was used to simulating the intracellular ROS concentration. High fitting accuracy was achieved by the model, demonstrating the effectiveness of the proposed modeling technique. Both the experimental data and the numerical results from the NARX model indicated that the ROS ratio of V. E. cells was always lower than that of *C. albicans*, demonstrating the safety of

the ABL therapy. Moreover, a key conclusion was that the 415nm wavelength blue light was the most effective wavelength, with the least damage to V. E. cells. This results was also confirmed in the previous chapter.

As a nonlinear modeling technique, the NARX model can simulate the ROS accumulation process better than the linear model. However, due to the nature of NARX, the form of the model was unclear and leads to some difficulties in using.

- Forth, an closed-form exact solution of MSO model was derived in Chapter 3. The contribution of this chapter has two aspects. First, the control theory was the first time tried applying on analysing the PDT model. After the pole-zero cancellation, it was found that this system is linear in an oxygen-rich environment. Thus, the MSO model can be reduced to a linear dynamic model. Second, simplified the MSO model to a closed-form analytical solution, which is much easier to use. This analytical solution estimate the singlet oxygen level with given light fluence and photon energy.

  The modeling technique was fitter to the experimental data on *C. albicans*. Results discoverer that the photon energy within the range of 2.7-2.8eV or 3.0-3.2eV (388-413nm or 443-459nm in wavelength) is more effective in generating singlet oxygen. This range perfectly cover the aforementioned treating wavelength.

7.2. LIMITATION AND FUTURE WORK

The proposed modeling techniques in this thesis offer opportunities for determining the light dosages in treating fungal infection diseases by PDT and ABL, especially those on the surface tissues of human body. However, all thees modeling approaches were developed based on the *in vitro* experimental data, the *in vivo* experiments were still lacked. One of the future work can be aimed at the *in vivo* verification of the proposed models.

In fact, some attempts have been tried by our team. Applying irradiation on an animal model was not difficult. Also, the survival rates of the pathogens can be studied by using the fluorescence fungi. The survival fungi can be observed directly by the fluorescence microscope. These methods have already been developed and successfully published bu other researchers. The main difficulty to verify our models is the measurement of the ROS and singlet oxygen concentration in the animal model. Because the short live time and unstable nature, ROS and singlet oxygen will react with oxygen, or transform into other species in a few seconds. Therefore, an instant measuring equipment in the animal model is requiring. The detector of the equipment needs to be small enough to be set up under the superficial skin of the animal. Due to the limitation of our lab facilities and funding, we did not make this equipment.

Another future aim can be clinical trials of applying ABL on treating microbial infection. However, the urgency of the alternative treatment of infection is not that much. People are still willing to use antibiotic or antimycotic drugs for their convenience and economy. This may become the reason that very few clinical trials were studied. In Chapter. 3, the spatial distribution of fluence rate and $^1O_2$ concentration under human epidermis was simulated. The simulation result was in accord with our knowledge, but future clinical trials on this topic would be a good supplement of the modeling approach.

ACKNOWLEDGEMENTS

After 4 years of postgraduate study, when I finished my thesis, I would like to express my sincere gratitude to the professors, colleagues, friends, and family who cared for and supported me by guiding and helping me.

Many great appreciations go to Prof. G.Q. Zhang. As my promotor, thanks for offering me the opportunity of being a Ph.D. candidate at TU Delft and supervising my work and thesis. I am sure my doctoral journey would not be so smooth without your kind help.

My thanks also go to Prof. J.F. Dong as my copromoter and daily supervisor. I dearly thank you for your guidance throughout the Ph.D. process. I wonder if I could have finished without your encouragement and guidance. Thanks for your effort spent supervising my study. I will also never forget the ideas you brought up to me; will never forget the time we spent debugging the program. I was complaining about your 'annoying' daily checking. However, when I finally come to this end and look back, I admit that checking makes me work much more efficiently. On the other hand, your academic attitude also impressed me a lot, which is a lesson that will prove very valuable for my career.

Then, thanks to my colleagues for helping, motivating, and caring for each other in our studies and life. Thanks for being by my side, Yun Zhao, Yunchu Zhang, and Qi Wang. In the evening after the weekly meeting, we always went out for good food, which brought me in a good mood. If I have to choose the happiest time in Suzhou, it will be the dinner time with your guys. Thanks will also go to Roel Stortelder, my dear dutch roommate, friend, and teammate, it was you to save me from the pressure, and save me in PUBG.

Appreciations must also be expressed to all the members of my Ph.D. defense committee, Prof. dr. ir. W.D. van Driel from Delft University of Technology, Prof. dr. R.A.M. Fouchier from Erasmus University Medical Center, Prof. dr. D.A. Pijnappels from Leiden University Medical Center, Prof. dr. Y.G. Tang from Suzhou Institute of Biomedical Engineering and Technology, Prof. dr. ir. W.L. IJzerman from Eindhoven University of Technology, and Prof. dr. ir. P.M. Sarro from Delft University of Technology. Thanks for spending time to review my thesis and giving me their valuable comments and feedback.

Finally, I would like to thank my parents, for their love, patience, support, and understanding, for so many years. When I decided to go abroad in 2015 to pursue my study, it was my parents to support me without any doubt. They certainly deserve to share the accomplishment of my Ph.D. study!

Tianfeng Wang
May 2022

A

APPENDIX

A.1. PSO ALGORITHM FOR KERNEL SELECTION

The PSO algorithm is an evolutionary algorithm that mimics the movements of the organisms in a bird flock. It solves a problem by having a population of candidate solutions, and moving these particles around in the search-space. The movements of the particles are guided by their own best known position in the search space as well as the best known position of the entire swarm.

First, all particles are dispersed uniformly. The movement, also known as velocity, is denoted as $\boldsymbol{V}$, whose initial value can be randomly chosen. Let $\boldsymbol{\Psi}_j, \boldsymbol{V}_j \in \mathbb{R}^S$ denotes the j-th particle and its corresponding velocity, where S denotes the dimension of $\boldsymbol{\Psi}_j$ and $\boldsymbol{V}_j$. The movements are adapted by the following formula.

$$\boldsymbol{V}_j = d_1 \boldsymbol{V}_j + d_2(\boldsymbol{\Psi}^* - \boldsymbol{\Psi}_j) + d_3(\boldsymbol{\Psi}^*_j - \boldsymbol{\Psi}_j), \tag{A.1}$$

where d_1, d_2 and d_3 are the weighting factors; $\boldsymbol{\Psi}^*$ and $\boldsymbol{\Psi}_j*$ are the global best particle and local best particle respectively. The value of $\boldsymbol{V}_j$ shall be hard bounded to the so-called "setting region" , i.e. $\boldsymbol{V}_j \in [\boldsymbol{V}_{min}, \boldsymbol{V}_{max}]$.

After each movement, the position of the j-th particle is updated by

$$\boldsymbol{\Psi}_{j+1} = \boldsymbol{\Psi}_j + \boldsymbol{V}_j, \tag{A.2}$$

which is then hard bounded to the search space, i.e. $\boldsymbol{\Psi}_j \in [\boldsymbol{\Psi}_{min}, \boldsymbol{V}_{max}]$. The search will stop when the cost function of $\boldsymbol{\Psi}^*$ is satisfied or the maximum generation is reached.

In this work, we propose to use PSO for the kernel selection in the NARX model. The objective of kernel selection is to select a subset of n_s ($n_s \ll n$) best kernels from the full set of n candidate kernels. In this specific NARX model, the I/O data are time-series data of the ROS concentrations which is a vector. To model each this process, the output dimension is $\ell = 1$, and hence y_k, θ_i take the scalar form. For simplicity of notations, we will especially consider the single output case. However, it shall be mentioned that the proposed method is not restricted to single output case. For the NARX kernel selection,

A

define a cost function as J_t, where t stands for the number of selected regressors. The initial cost is denoted as $J_0 = \boldsymbol{y}^T \boldsymbol{y}$. According to Eq. (8) one can write

$$J_0 = \boldsymbol{\theta}^T \boldsymbol{P}^T \boldsymbol{P} \boldsymbol{\theta} = \Sigma_{i=1}^{n} \boldsymbol{p}_i^T \boldsymbol{p}_i \theta_i^2.$$

The idea is to search all the columns in $\boldsymbol{P}$, and find the one that reduces the cost value most, if being removed from J.

$$J_i = J_{i-1} - \boldsymbol{p}_i^T \boldsymbol{p}_i \theta_i^2, \tag{A.3}$$

where $\boldsymbol{p}_i$ is chosen from $\boldsymbol{P}$. This procedure can be terminated if

$$J_i < \epsilon \ or \ i < T \tag{A.4}$$

is satisfied, where ϵ is a chosen positive scalar; and "T" is the maximum epoch, represents that T regressors are chosen.

The PSO algorithm is proposed to choose the orthogonal columns in Eq. (A.3). In PSO, the decision variables are regarded as particles. These particles move around in the search space.

The parameters in the NARX model to be optimized are the kernel center $\boldsymbol{x}_i$ and the kernel parameters $a_i, b_i, \cdots$. Thus, for kernel selection, the particle shall include the index of the center and its corresponding variance; i.e., $\boldsymbol{\Psi}_j = [i_j, a_{i_j}, b_{i_j}, \cdots]^T$, where i_j stands for the index of the kernel center that is included in the j-th particle while choosing $\boldsymbol{p}_i$. Once these kernel parameters are fixed, the corresponding $\boldsymbol{p}_i$ and θ_i can be calculated by the standard Gram-Schmidt procedure; and the cost function can be calculated by Eq. (A.3).

A.2. NARX MODEL IDENTIFICATION ALGORITHM

Outer loop (search the i-th regressor $\boldsymbol{g}_i$, $i = 1, 2, \cdots, T$)
Initialize the value of population size denoted as Q; the weighting factors d_1, d_2 and d_3; the searching space $\boldsymbol{\Psi}_{min}, \boldsymbol{\Psi}_{max}$; and the movement setting region $\boldsymbol{V}_{min}, \boldsymbol{V}_{max}$.
for(i=1; i ≤ T; i=i+1)

1. Initialize the population $\boldsymbol{\Psi}$ randomly in the searching space and randomly generate an initial velocity matrix $\boldsymbol{V}$ inside the setting region.

2. Calculate the cost for all particles:

   for(j=1; j ≤ Q; j=j+1)

   (a) Choose the kernel index i_j from the particle $\boldsymbol{\Psi}_j = [i_j, a_{i_j}, b_{i_j}, \cdots]^T$. Compute the regressor vector $\boldsymbol{g}_{i_j} = [g_{i_j}(\boldsymbol{x}_1), \cdots, g_{i_j}(\boldsymbol{x}_N)]^T$ by the kernel functions on Table (2), as a candidate while selecting the i-th regressor. Then orthogonalize it by the standard Gram-Schmidt procedure:

(b) if $i = 1$

$$\begin{aligned} \boldsymbol{u}_j &= \boldsymbol{g}_{i_j} \\ \boldsymbol{p}_{i_j} &= \frac{\boldsymbol{u}_j}{\|\boldsymbol{u}_j\|_2} \end{aligned} \tag{A.5}$$

else

$$\begin{aligned} \boldsymbol{u}_j &= \boldsymbol{g}_{i_j} - \Sigma_{r=1}^{i-1}(\boldsymbol{g}_{i_j}^T \cdot \boldsymbol{p}_r) \cdot \boldsymbol{p}_r \\ \boldsymbol{p}_{i_j} &= \frac{\boldsymbol{u}_j}{\|\boldsymbol{u}_j\|_2} \end{aligned} \tag{A.6}$$

(c) Calculate the corresponding θ_{i_j} and the cost function in terms of $\boldsymbol{p}_{i_j}, \theta_{i_j}$ by

$$\theta_{i_j} = \frac{{\boldsymbol{p}_{i_j}}^T y}{{\boldsymbol{p}_{i_j}}^T \boldsymbol{p}_{i_j}} \tag{A.7}$$

$$J_{i_j} = J_{i_{j-1}} - {\boldsymbol{p}_{i_j}}^T \boldsymbol{p}_{i_j} {\theta_{i_j}}^2 \tag{A.8}$$

where J_{i_j} is the cost after choose $\boldsymbol{\Psi}_j = [i_j, a_{i_j}, b_{i_j}, \cdots]^T$ as the i-th regressor.

end for

3. Find the best particle, as the one corresponding to the minimal J_{i_j}, i.e.

$$j* = \min_j \{J_{i_1}, J_{i_2}, \cdots, J_{i_Q}\}.$$

And set the global best as $\Psi^* = \Psi_{j*}$.

4. Initialize the local best particles as

$$\Psi_{j*} = \Psi_j, j = 1, 2, \cdots, Q$$

The PSO inner loop:

for(k=1; k ≤ K; k=k+1)

for($j = 1; j \le Q, j = j + 1$)

(a) Calculate $\boldsymbol{V}_j$ by (A.1) and restrict its value into the setting region.

(b) Update $\boldsymbol{\Psi}_j$ by (A.2) and restrict its value into the search space.

(c) Calculate the cost function value J_{i_j} as in Eq. (A.8) for the new particle.
Update the global best particle $\boldsymbol{\Psi}^*$ and the local best particles $\boldsymbol{\Psi}_i^*$ as follows.
If $\boldsymbol{\Psi}_j$ is better than $\boldsymbol{\Psi}^*$, set $\boldsymbol{\Psi}^* = \boldsymbol{\Psi}_j$.
If $\boldsymbol{\Psi}_j$ is better than $\boldsymbol{\Psi}_j^*$, set $\boldsymbol{\Psi}_j^* = \boldsymbol{\Psi}_j$.

end for

Stop the inner loop when the maximum epoch is reached or $J_{i_j}^* < \epsilon$.

A

end for

End of inner loop

The best particle find in this epoch is $\boldsymbol{\Psi}^*$. Set $J_i = J_{i_{j*}}$, $\boldsymbol{g}_i = \boldsymbol{g}_{i_{j*}}$, $\boldsymbol{p}_i = \boldsymbol{p}_{i_{j*}}$ and $\boldsymbol{\theta}_i = \boldsymbol{\theta}_{i_{j*}}$. Remove the selected index $j*$ from the candidates of the regressor indices.

end for

End of outer loop

A.3. THE RATE EQUATIONS OF EGFR SIGNALING NETWORK

The rate equations of EGFR signaling network are listed below, where C_* denoted the concentration of protein "$*$"; K and k with subscript are respectively the Michaelis constant and the kinetic constant.

$$
\begin{aligned}
v_1 &= k_1 C_R C_{EGF} - k_{-1} C_{R_a} \\
v_2 &= k_2 C_{R_a} C_{R_a} - k_{-2} C_{R_2} \\
v_3 &= k_3 C_{R_2} - k_{-3} C_{RP} \\
v_4 &= V_4 C_{RP} / (K_4 + C_{RP}) \\
v_5 &= k_5 C_{RP} C_{PLC\gamma} - k_{-5} C_{RPL} \\
v_6 &= k_6 C_{R-PL} - k_{-6} C_{R-PLP} \\
v_7 &= k_7 C_{R-PLP} - k_{-7} C_{RP} C_{PLC\gamma P} \\
v_8 &= V_8 C_{PLC\gamma P} / (K_8 + C_{PLC\gamma P}) \\
v_9 &= k_9 C_{RP} C_{Grb} - k_{-9} C_{R-G} \\
v_{10} &= k_{10} C_{R-G} C_{SOS} - k_{-10} C_{R-G-S} \\
v_{11} &= k_{11} C_{R-G-S} - k_{-11} C_{RP} C_{G-S} \\
v_{12} &= k_{12} C_{G-S} - k_{-12} C_{Grb} C_{SOS} \\
v_{13} &= k_{13} C_{RP} C_{Shc} - k_{-13} C_{R-Sh} \\
v_{14} &= k_{14} C_{R-Sh} - k_{-14} C_{R-ShP} \\
v_{15} &= k_{15} C_{R-ShP} - k_{-15} C_{RP} C_{ShP} \\
v_{16} &= V_{16} C_{ShP} / (K_{16} + C_{ShP}) \\
v_{17} &= k_{17} C_{R-ShP} C_{Grb} - k_{-15} C_{R-Sh-G} \\
v_{18} &= k_{18} C_{R-Sh-G} - k_{-18} C_{RP} C_{ShG} \\
v_{19} &= k_{19} C_{R-Sh-G} C_{SOS} - k_{-19} C_{R-Sh-G-S} \\
v_{20} &= k_{20} C_{R-Sh-G-S} - k_{-20} C_{Sh-G-S} C_{RP} \\
v_{21} &= k_{21} C_{ShP} C_{Grb} - k_{-21} C_{Sh-G} \\
v_{22} &= k_{22} C_{Sh-G} C_{SOS} - k_{-22} C_{Sh-G-S} \\
v_{23} &= k_{23} C_{Sh-G-S} - k_{-23} C_{ShP} C_{G-S} \\
v_{24} &= k_{24} C_{R-ShP} C_{G-S} - k_{-24} C_{R-Sh-G-S} \\
v_{25} &= k_{24} C_{PLC\gamma P} - k_{-25} C_{PLC\gamma P-I}
\end{aligned}
$$

The kinetic equations are derived according the EGFR network in Fig. 5.1.

$$
\begin{aligned}
\frac{d[C_{EGF}]}{dt} &= -v_1 \\
\frac{d[C_R]}{dt} &= -v_1 \\
\frac{d[C_{R_a}]}{dt} &= v_1 - 2v_2 \\
\frac{d[C_{R_2}]}{dt} &= v_2 + v_4 - v_3 \\
\frac{d[C_{RP}]}{dt} &= v_3 + v_7 + v_{11} + v_{15} + v_{18} + v_{20} - v_4 - v_5 - v_9 - v_{13} \\
\frac{d[C_{R-PL}]}{dt} &= v_5 - v_6 \\
\frac{d[C_{R-PLP}]}{dt} &= v_6 - v_7 \\
\frac{d[C_{R-G}]}{dt} &= v_9 - v_{10} \\
\frac{d[C_{R-G-S}]}{dt} &= v_{10} - v_{11} \\
\frac{d[C_{R-Sh}]}{dt} &= v_{13} - v_{14} \\
\frac{d[C_{R-ShP}]}{dt} &= v_{14} - v_{24} - v_{15} - v_{17} \\
\frac{d[C_{R-Sh-G}]}{dt} &= v_{17} - v_{18} - v_{19} \\
\frac{d[C_{R-Sh-G-S}]}{dt} &= v_{19} - v_{20} + v_{21} \\
\frac{d[C_{G-S}]}{dt} &= v_{11} + v_{23} - v_{12} + v_{24} \\
\frac{d[C_{ShP}]}{dt} &= v_{15} + v_{23} - v_{21} + v_{16} \\
\frac{d[C_{Sh-G}]}{dt} &= v_{18} + v_{21} - v_{22} \\
\frac{d[C_{PLC\gamma}]}{dt} &= v_8 - v_5
\end{aligned}
$$

A

Table A.1: The index of the all 23 proteins

1	EGF	5	RP	9	R-G-S	13	R-Sh-G-S	17	PLCγ	21	Shc
2	R	6	R-PL	10	R-Sh	14	G-S	18	PLCγP	22	SOS
3	Ra	7	R-PLP	11	R-Shp	15	Shp	19	PLCγP-I	23	Sh-G-S
4	R2	8	RG	12	R-Sh-G	16	ShG	20	Grb		

$$
\begin{aligned}
\frac{d[C_{PLC\gamma P}]}{dt} &= v_7 - v_5 - v_{25} \\
\frac{d[C_{PLC\gamma P-I}]}{dt} &= v_{25} \\
\frac{d[C_{Grb}]}{dt} &= v_{12} - v_9 - v_{17} - v_{21} \\
\frac{d[C_{Shc}]}{dt} &= v_{16} - v_{11} \\
\frac{d[C_{SOS}]}{dt} &= v_{12} - v_{10} - v_{19} - v_{22} \\
\frac{d[C_{Sh-G-S}]}{dt} &= v_{20} + v_{22} - v_{23}
\end{aligned}
$$

CURRICULUM VITÆ

Tianfeng WANG

1992 Born in Changchun, Jilin, China.

EDUCATION

2011 Bachelor of Engineering in Mechanical Engineering and Manufacturing and Automation, University of Science and Technology of China, Hefei, Anhui, China
Thesis: Shaping Micro Droplet through Designing the Microstructure of Substrate

2015 Master of Science in Mechanical Engineering, State University of New York at Buffalo, Buffalo, NY, US
Project: Quadrotor Optimization by Heuristic Algorithms

2017 Ph.D. in Department of Microelectronics, Electrical Engineering & Mathematics & Computer Science, Delft University of Technology, Delft, the Netherlands
Thesis: Light Dosage Optimization by Data-driven and Dynamic modeling *in blue light therapies*
Promotor: Prof. dr. Guoqi Zhang

LIST OF PUBLICATIONS

6. **Tianfeng. Wang**, Jianfei Dong, Huancai Yin, Guoqi Zhang, *Blue light therapy to treat candida vaginitis with comparisons of three wavelengths: an in vitro study*, Lasers in medical science **35**, 6 (2020).

5. Jianfei Dong, **Tianfeng Wang**, *Data driven modeling of the reactive oxygen species stimulated by photon energy in light therapies*, IEEE Access **8**, 891 (2020).

4. **Tianfeng. Wang**, Jianfei Dong, Guoqi Zhang, *Analyzing efficacy and safety of anti-fungal blue light therapy via kernel-based modeling the reactive oxygen species induced by light*, IEEE Transactions on Biomedical Engineering.

3. **Tianfeng. Wang**, Jianfei Dong, Guoqi Zhang, *Analysis and data-based modeling of the photochemical reaction dynamics of the induced singlet oxygen in light therapies*, IEEE Transactions on Biomedical Engineering.

2. **Tianfeng Wang,** Jianfei Dong, Guoqi Zhang, *Applying LEDs as Therapeutic Light Sources for Anti-microbial Treatment: An Experimental Study*, 2019 16th China International Forum on Solid State Lighting & 2019 International Forum on Wide Bandgap Semiconductors China (SSLChina: IFWS).

1. Jianfei Dong, **Tianfeng Wang,** Yun Zhao, Yunchu Zhang, *Acquiring and mining the data from light-cell interaction experiments*, In 2020 17th China International Forum on Solid State Lighting & 2020 International Forum on Wide Bandgap Semiconductors China (SSLChina: IFWS) .